

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
23. Oktober 2003 (23.10.2003)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 03/087386 A2

- (51) Internationale Patentklassifikation<sup>7</sup>: C12P 13/12, (74) Anwalt: SCHWEIGER Georg; Reistötter, Kinzebach & Partner (GbR), Sternwartstrasse 4, 81679 München (DE).  
C12N 9/10, 1/15, 15/63
- (21) Internationales Aktenzeichen: PCT/EP03/04010 (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (22) Internationales Anmeldedatum:  
16. April 2003 (16.04.2003)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:  
102 17 058.4 17. April 2002 (17.04.2002) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; 67056 Ludwigshafen (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): KRÖGER, Burkhard [DE/DE]; Im Waldhof 1, 67117 Limburgerhof (DE). ZELDER, Oskar [DE/DE]; Franz-Stützel-Str. 8, 67346 Speyer (DE). KLOPPROGGE, Corinna [DE/DE]; Diemersteinstrasse 3, 67065 Ludwigshafen (DE). SCHRÖDER, Hartwig [DE/DE]; Benzstr. 4, 69226 Nussloch (DE). HÄFNER, Stefan [DE/DE]; Luitpoldstr. 11, 67063 Ludwigshafen (DE).
- (84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Veröffentlicht:  
— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts
- Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

WO 03/087386 A2

(54) Title: METHODS FOR PRODUCING SULPHUROUS FINE CHEMICALS

(54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG VON SCHWEFELHALTIGEN FEINCHEMIKALIEN

(57) Abstract: The invention relates to methods for producing sulphurous fine chemicals, in particular L-methionine, by fermentation, using bacteria, in which a nucleotide sequence that codes for a methionine synthase (metF) gene is expressed.

(57) Zusammenfassung: Die Erfindung betrifft Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für ein Methionin-Synthase (metF)-Gen kodierende Nukleotidsequenzen exprimiert wird.

## **Verfahren zur Herstellung von schwefelhaltigen Feinchemikalien**

### **Beschreibung**

Gegenstand der Erfindung ist ein Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für ein Methionin-Synthase (methH)-Gen kodierende Nukleotidsequenz exprimiert wird.

### **Stand der Technik**

Schwefelhaltige Feinchemikalien, wie zum Beispiel Methionin, Homocystein, S-Adenosyl-Methionin, Glutathion, Cystein, Biotin, Thiamin, Liponsäure werden über natürliche Stoffwechselprozesse in Zellen hergestellt und werden in vielen Industriezweigen verwendet, einschließlich der Nahrungsmittel-, Futtermittel-, Kosmetik- und pharmazeutischen Industrie. Diese Substanzen, die zusammen als "schwefelhaltige Feinchemikalien" bezeichnet werden, umfassen organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Vitamine und Cofaktoren. Ihre Produktion erfolgt am zweckmäßigsten im Großmaßstab mittels Anzucht von Bakterien, die entwickelt wurden, um große Mengen der jeweils gewünschten Substanz zu produzieren und sezernieren. Für diesen Zweck besonders geeignete Organismen sind coryneforme Bakterien, gram-positive nicht-pathogene Bakterien.

Es ist bekannt, dass Aminosäuren durch Fermentation von Stämmen coryneformer Bakterien, insbesondere *Corynebacterium glutamicum*, hergestellt werden. Wegen der großen Bedeutung wird ständig an der Verbesserung der Herstellverfahren gearbeitet. Verfahrensverbesserungen können fermentationstechnische Maßnahmen, wie zum Beispiel Rührung und Versorgung mit Sauerstoff, oder die Zusammensetzung der Nährmedien, wie zum Beispiel die Zuckerkonzentration während der Fermentation, oder die Aufarbeitung zum Produkt, beispielsweise durch Ionenaustauschchromatographie, oder die intrinsischen Leistungseigenschaften des Mikroorganismus selbst betreffen.

Über Stammselektion sind eine Reihe von Mutantenstämmen entwickelt worden, die ein Sortiment wünschenswerter Verbindungen aus der Reihe der schwefelhaltigen Feinchemikalien produzieren. Zur Verbesserung der Leistungseigenschaften dieser Mikroorganismen hinsichtlich der Produktion eines bestimmten Moleküls werden Methoden der Mutagenese, Selektion und Mutantenauswahl angewendet. Dies ist jedoch ein zeitaufwendiges und schwieriges Verfahren. Auf

diese Weise erhält man z.B. Stämme, die resistent gegen Antimetabolite oder Hemmstoffe, wie z. B. die Methionin-Analoga  $\alpha$ -Methyl-Methionin, Ethionin, Norleucin, N-Acetylnorleucin, S-Trifluoromethylhomocystein, 2-Amino-5-heptenoisäure, Seleno-Methionin, Methioninsulfoximin, Methoxin, 1-Aminocyclopentan-Carboxylsäure oder auxotroph für regulatorisch bedeutsame Metabolite sind und schwefelhaltige Feinchemikalien, wie z. B. L-Methionin, produzieren.

Seit einigen Jahren werden ebenfalls Methoden der rekombinanten DNA-Technik zur Stammverbesserung von L-Aminosäure produzierender Stämme von Corynebacterium eingesetzt, indem man einzelne Aminosäure-Biosynthesegene amplifiziert und die Auswirkung auf die Aminosäure-Produktion untersucht.

Die WO-A-02/10209 beschreibt ein Verfahren zur fermentativen Herstellung von L-Methionin unter Verwendung L-Methionin produzierender coryneformer Bakterien, worin wenigstens das methH-Gen überexprimiert ist, sowie die kodierende methH-Sequenz aus C. glutamicum ATCC 13032.

#### Kurze Beschreibung der Erfindung

Der Erfindung lag die Aufgabe zugrunde, ein neues Verfahren zur verbesserten fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, bereitzustellen.

Gelöst wird obige Aufgabe durch Bereitstellung eines Verfahrens zur fermentativen Herstellung einer schwefelhaltigen Feinchemikalie, umfassend die Expression einer heterologen Nukleotidsequenz, welche für ein Protein mit methH-Aktivität kodiert, in einem coryneformen Bakterium.

Ein erster Gegenstand der Erfindung ist ein Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen Feinchemikalie, welches folgende Schritte umfasst:

- a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie produzierenden coryneformen Bakterienkultur, wobei in den coryneformen Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird, welche für ein Protein mit Methionin-Synthase (methH) –Aktivität kodiert;
- b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den Zellen der Bakterien, und

## 3

- c) Isolieren der schwefelhaltigen Feinchemikalie, welche vorzugsweise L-Methionin umfasst.

- 5 Vorzugsweise besitzt obige heterologe methH-kodierende Nukleotidsequenz zur methH-kodierenden Sequenz aus *Corynebacterium glutamicum* ATCC 13032 eine Sequenzhomologie vom weniger als 70% aufweist. Die methH-kodierende Sequenz ist vorzugsweise aus einem der folgenden Organismen von Liste I abgeleitet:

Liste I

10

<i>Streptomyces coelicolor</i>	ATCC 10147
<i>Anabaena</i> sp.	ATCC 27892
<i>Synechocystis</i> sp.	ATCC 27184
<i>Prochlorococcus marinus</i>	PCC 7118
<i>Thermus thermophilus</i>	ATCC 27634
<i>Bacillus halodurans</i>	ATCC 21591
<i>Bacillus stearothermophilus</i>	ATCC 12980
<i>Vibrio cholerae</i>	ATCC 39315
<i>Sinorhizobium meliloti</i>	ATCC 4399
<i>Escherichia coli</i> K12	ATCC 55151
<i>Salmonella typhimurium</i>	ATCC 15277
<i>Salmonella typhi</i>	ATCC 12839
<i>Pseudomonas fluorescens</i>	ATCC 13525
<i>Pseudomonas aeruginosa</i>	ATCC 17933
<i>Nitrosomonas europaea</i>	ATCC 19718
<i>Bordetella pertussis</i>	ATCC 9797
<i>Clorobium tepidum</i>	ATCC 49652
<i>Deinococcus radiodurans</i>	ATCC 13939
<i>Clostridium acetobutylicum</i>	ATCC 824
<i>Caulobacter crescentus</i>	ATCC 19089
<i>Homo sapiens</i>	
<i>Vibrio fischeri</i>	ATCC 33715
<i>Agrobacterium tumefaciens</i> str. C58 (Cereon)	ATCC 33970
<i>Ralstonia solanacearum</i>	ATCC 25237

ATCC: American Type Culture Collection, Rockville, MD, USA

PCC: Pasteur Culture Collection of Cyanobacteria. Paris Frankreich

- 15 Die erfindungsgemäß eingesetzte methH-kodierende Sequenz umfasst vorzugsweise eine kodierende Sequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35,



37, 39, 41, 43, 45, 47, 49 und 51 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit methH-Aktivität kodiert.

5 Die erfindungsgemäß eingesetzte methH-kodierende Sequenz kodiert außerdem vorzugsweise für ein Protein mit methH-Aktivität, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 und 52 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit methH-Aktivität steht, umfasst.

10 Die kodierende methH-Sequenz ist vorzugsweise eine in coryneformen Bakterien replizierbare oder eine stabil in das Chromosom integrierte DNA oder eine RNA.

Gemäß einer bevorzugten Ausführungsform wird das erfindungsgemäße Verfahren durchgeführt, indem man

15

- a) einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden methH-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
- b) einen Stamm einsetzt, in dem die kodierende methH-Sequenz in das Chromosom des Bakteriums integriert wurde.

20

Es ist weiterhin bevorzugt, die kodierende methH-Sequenz für die Fermentation zu überexprimieren.

25

Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie oder eines damit assoziierten Biosynthese- oder sonstigen Stoffwechselweges, verstärkt ist; und / oder

in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet sind, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.

30

Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie durch Stoffwechselmetabolite in seiner Aktivität nicht in unerwünschter Weise beeinflusst wird.

Gemäß einer weiteren Ausführungsform des erfindungsgemäßen Verfahrens werden deshalb coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

5

- a) dem für eine Aspartatkinase kodierenden Gen *lysC*,
  - b) dem für eine Aspartat-Semialdehyd-Dehydrogenase kodierenden Gen *asd*
  - c) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen *gap*,
  - d) dem für die 3-Phosphoglycerat Kinase kodierenden Gen *pgk*,
  - 10 e) dem für die Pyruvat Carboxylase kodierenden Gen *pyc*,
  - f) dem für die Triosephosphat Isomerase kodierenden Gen *tpi*,
  - g) dem für die Homoserin O-Acetyltransferase kodierenden Gen *metA*,
  - h) dem für die Cystathionin-gamma-Synthase kodierenden Gen *metB*,
  - i) dem für die Cystathionin-gamma-Lyase kodierenden Gen *metC*,
  - 15 j) dem für die Serin-Hydroxymethyltransferase kodierenden Gen *glyA*,
  - k) dem für die O-Acetylhomoserin-Sulphydrylase kodierenden Gen *metY*,
  - l) dem für die Methylen-Tetrahydrofolat-Reduktase kodierenden Gen, *metF*
  - m) dem für die Phosphoserin-Aminotransferase kodierenden Gen *serC*
  - n) dem für die Phosphoserin-Phosphatase kodierenden Gen *serB*,
  - 20 o) dem für die Serine Acetyl-Transferase kodierenden Gen *cysE*,
  - p) dem für die Homoserin-Dehydrogenase kodierenden Gen *hom*,
- überexprimiert ist.

25

Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene ausgewählt unter Genen der oben genannten Gruppe a) bis p) mutiert ist, insbesondere so dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden und dass insbesondere die erfindungsgemäße Produktion der Feinchemikalie nicht beeinträchtigt wird. Durch die Mutation kann

30 das Protein auch eine höhere Aktivität (Sunstratumsatz) und/oder Sunstratspezifität besitzen und damit die Produktion der gewünschten Feinchemikalie fördern.

Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- q) dem für die Homoserine-Kinase kodierenden Gen thrB,
- r) dem für die Threonin Dehydratase kodierenden Gen ilvA,
- 5 s) dem für die Threonin Synthase kodierenden Gen thrC
- t) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen ddh
- u) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen pck,
- v) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen pgj,
- w) dem für die Pyruvat-Oxidase kodierenden Gen poxB,
- 10 x) dem für die Dihydrodipicolinat Synthase kodierenden Gen dapA,
- y) dem für die Dihydrodipicolinat Reduktase kodierenden Gen dapB; oder
- z) dem für die Diaminopicolinat Decarboxylase kodierenden Gen lysA

abschwächt ist, insbesondere durch Verringerung der Expressionsrate des korrespondierenden Gens, oder durch Expression eines Proteins mit geringerer Aktivität (Substratumsatz).

15

Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene der obigen Gruppen q) bis z) mutiert ist, so dass die enzymatische Aktivität des korrespondierenden Proteins teilweise oder vollständig verringert wird.

20

Vorzugsweise werden in dem erfindungsgemäßen Verfahren Mikroorganismen der Art *Corynebacterium glutamicum* eingesetzt.

25

In einer weiteren Ausführungsform des Verfahrens werden solche Mikroorganismen eingesetzt, die Resistenz gegenüber wenigstens einen Methionin-Biosynthesehemmer aufweisen. Solche Hemmer sind, ohne darauf beschränkt zu sein, Methionin-Analoga, wie  $\alpha$ -Methyl-Methionin, Ethionin, Norleucin, N-Acetylnorleucin, S-Trifluoromethylhomocystein, 2-Amino-5-heptenoisäure, Seleno-Methionin, Methioninsulfoximin, Methoxin, und 1-Aminocyclopentan-Carboxylsäure.

30

Ein weiterer Gegenstand der Erfindung betrifft ein Verfahren zur Herstellung eines L-Methionin-haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst

## 7

- 5
- a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
  - b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
  - c) Entfernung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
  - d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.

10 Gegenstand der Erfindung sind ebenfalls die erstmalig aus obigen Mikroorganismen isolierten kodierenden methH-Sequenzen, die davon kodierten Methionin-Synthasen sowie die funktionalen Homologen dieser Polynukleotide bzw. Proteine.

15 Gegenstand der Erfindung sind insbesondere auch die zur Durchführung obiger Verfahren notwendigen Expressionskonstrukte und Mikroorganismen.

15 Weitere Gegenstände der Erfindung sind somit insbesondere:

20 - das Plasmid pCIS lysC thr311ile, kodierend für lysC thr311ile oder ein funktionales Äquivalent davon, d.h. eine lysC-Mutante mit vergleichbarer, gegenüber dem Wildtyp erhöhter Aspartatkinase-Aktivität;

25 - ein Wirtsorganismus transformiert mit dem Plasmid pCIS lysC thr311ile, insbesondere ausgewählt unter Mikroorganismen der Gattung Corynebacterium, insbesondere der Art C. glutamicum, wie der transformierte Stamm LU1479 lysC 311ile;

25 - das Plasmid pC Phsdh methH\_Sc, kodierend für methH aus Streptomyces coelicolor;

30 - ein Wirtsorganismus gemäß obiger Definition, transformiert mit einem Plasmid, kodierend für exogenes methH; insbesondere transformiert mit dem Plasmid pC Phsdh methH\_Sc;

30 - ein Wirtsorganismus gemäß obiger Definition mit Resistenz gegen wenigstens einen Methionin-Biosynthesehemmstoff, wie der transformierte Stamm LU1479 lysC 311ile ET-16, gege-

benenfalls transformiert mit einer exogenen kodierenden meth Sequenz, wie der transformierte Stamm LU1479 lysC 311ile ET-16 pC Phsdh meth\_Sc.

#### Detaillierte Beschreibung der Erfindung

5

##### a) Allgemeine Begriffe

Als Proteine mit der biologischen Aktivität der Methionine-Synthase, kurz auch als methH genannt (systematische Bezeichnung: 5-Methyltetrahydrofolat-Homocystein S-Methyltransferase ;EC 10 2.1.1.13), werden solche Proteine bezeichnet, die in der Lage sind Homocystein unter Verwendung der Cofaktoren 5-Methyltetrahydrofolat (MTHF), Cobalamin (Vitamin B12) und S-Adenosyl-Methionin zu Methionin und Tetrahydrofolat umzusetzen. Während der Cofaktor 5-Methyltetrahydrofolat stöchiometrisch in die Reaktion mit eingeht (1mol MTHF/1Mol Methionin gebildet) wird, wie in der Literatur beschrieben, S-Adenosyl-Methionin substöchiometrisch umgesetzt. Cobalamin hingegen ist katalytisch an der Umsetzung beteiligt. Dem Fachmann sind 15 weitere Details des methH-Proteins bekannt. (Banerjee R.V., Matthews R.G. FASEB J. 4:1450-1459, 1990, Ludwig ML. Matthews RG. Annual Review of Biochemistry. 66:269-313, 1997, Drennan CL. Matthews RG. Ludwig ML. Current Opinion in Structural Biology. 4:919-29, 1994). Der Fachmann unterscheidet die Aktivität der Cobalamin-abhängigen 5-Methyltetrahydrofolat-Homocystein S-Methyltransferase von der der Cobalamin-unabhängigen 5-Methyltetrahydro- 20 Pteroyltriglutamat-Homocystein S-Methyltransferase (EC 2.1.1.14) auch metE genannt. Der Fachmann kann die enzymatische Aktivität von methH durch Enzymtests nachweisen, Vorschriften dafür können sein: Jarrett JT. Goulding CW. Fluhr K. Huang S. Matthews RG. Methods in Enzymology. 281:196-213, 1997.

25

Im Rahmen der vorliegenden Erfindung umfasst der Begriff „schwefelhaltige Feinchemikalie“ jegliche chemische Verbindung, die wenigstens ein Schwefelatom kovalent gebunden enthält und durch ein erfindungsgemäßes Fermentationsverfahrens zugänglich ist. Nichtlimitierende Beispiele dafür sind Methionin, Homocystein, S-Adenosyl-Methionin, insbesondere Methio- 30 nin, und S-Adenosyl-Methionin.

Im Rahmen der vorliegenden Erfindung umfassen die Begriffe L-Methionin, Methionin, Homocystein und S-Adenosylmethionin auch die korrespondierenden Salze, wie z. B. Methionin-Hydrochlorid oder Methionin-Sulfat.

- 5 "Polynukleotide" bezeichnet im allgemeinen Polyribonukleotide (RNA) und Polydeoxyribonukleotide (DNA), wobei es sich um nicht modifizierte RNA oder DNA oder modifizierte RNA oder DNA handeln kann.

- 10 Unter "Polypeptiden" versteht man erfindungsgemäß Peptide oder Proteine, die zwei oder mehr über Peptidbindungen verbundene Aminosäuren enthalten.

- Der Begriff „Stoffwechselmetabolit“ bezeichnet chemische Verbindungen, die im Stoffwechsel von Organismen als Zwischen- oder auch Endprodukte vorkommen und die neben ihrer Eigenschaft als chemische Bausteine auch modulierende Wirkung auf Enzyme und ihre katalytische
- 15 Aktivität haben können. Dabei ist aus der Literatur bekannt, dass solche Stoffwechselmetabolite sowohl hemmend als auch stimulierend auf die Aktivität von Enzymen wirken können (Biochemistry, Stryer, Lubert, 1995 W. H. Freeman & Company, New York, New York.). In der Literatur ist auch beschrieben, dass es möglich ist durch Maßnahmen wie Mutation der genomischen DNA durch UV-Strahlung, ionisierender Strahlung oder mutagene Substanzen und nachfolgen-
- 20 der Selektion auf bestimmte Phänotypen in Organismen solche Enzyme zu produzieren, in denen die Beeinflussung durch Stoffwechselmetabolite verändert wurde (Sahm H. Eggeling L. de Graaf AA. Biological Chemistry 381(9-10):899-910, 2000; Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94). Diese veränderten Eigenschaften können auch durch gezielte Maßnahmen erreicht werden. Dabei ist dem Fachmann bekannt, in Genen für
- 25 Enzyme auch gezielt bestimmte Nukleotide der für das Protein kodierenden DNA so zu verändern, dass das aus der exprimierten DNA-Sequenz resultierende Protein bestimmte neue Eigenschaften aufweist, so zum Beispiel, dass die modulierende Wirkung von Stoffwechselmetaboliten gegenüber dem nicht veränderten Protein verändert ist
- 30 Enzyme können derart in ihrer Aktivität beeinflusst werden, dass es zu einer Verringerung der Reaktionsgeschwindigkeit, oder zu einer Veränderung der Affinität gegenüber dem Substrat oder zu einer Änderung der Reaktionsgeschwindigkeiten kommt

Die Begriffe "exprimieren" bzw. "Verstärkung" oder „Überexpression“ beschreiben im Kontext der Erfindung die Produktion bzw. Erhöhung der intrazellulären Aktivität eines oder mehrerer Enzyme in einem Mikroorganismus, die durch die entsprechende DNA kodiert werden. Dazu kann man beispielsweise ein Gen in einen Organismus einbringen, ein vorhandenes Gen durch ein  
5 anderes Gen ersetzen, die Kopienzahl des Gens bzw. der Gene erhöhen, einen starken Promotor verwenden oder ein Gen verwenden, das für ein entsprechendes Enzym mit einer hohen Aktivität kodiert und man kann gegebenenfalls diese Maßnahmen kombinieren.

b) Erfindungsgemäße methH-Proteine

10

Erfindungsgemäß mit umfasst sind ebenfalls „funktionale Äquivalente“ der konkret offenbarten methH-Enzyme aus Organismen obiger Liste I.

15 „Funktionale Äquivalente“ oder Analoga der konkret offenbarten Polypeptide sind im Rahmen der vorliegenden Erfindung davon verschiedene Polypeptide, welche weiterhin die gewünschte biologische Aktivität, wie z.B. Substratspezifität, besitzen.

Unter "funktionalen Äquivalenten" versteht man erfindungsgemäß insbesondere Mutanten, welche in wenigstens einer der oben genannten Sequenzpositionen eine andere als die konkret  
20 genannte Aminosäure aufweisen aber trotzdem eine der oben genannten biologischen Aktivitäten besitzen. "Funktionale Äquivalente" umfassen somit die durch eine oder mehrere Aminosäure-Additionen, -Substitutionen, -Deletionen und/oder -Inversionen erhältlichen Mutanten, wobei die genannten Veränderungen in jeglicher Sequenzposition auftreten können, solange sie zu einer Mutante mit dem erfindungsgemäßen Eigenschaftsprofil führen. Funktionale Äquivalenz ist  
25 insbesondere auch dann gegeben, wenn die Reaktivitätsmuster zwischen Mutante und unverändertem Polypeptid qualitativ übereinstimmen, d.h. beispielsweise gleiche Substrate mit unterschiedlicher Geschwindigkeit umgesetzt werden.

30 "Funktionale Äquivalente" umfassen natürlich auch Polypeptide welche aus anderen Organismen zugänglich sind, sowie natürlich vorkommende Varianten. Beispielsweise lassen sich durch Sequenzvergleich Bereiche homologer Sequenzregionen festlegen und in Anlehnung an die konkreten Vorgaben der Erfindung äquivalente Enzyme ermitteln.

„Funktionale Äquivalente“ umfassen ebenfalls Fragmente, vorzugsweise einzelne Domänen oder Sequenzmotive, der erfindungsgemäßen Polypeptide, welche z.B. die gewünschte biologische Funktion aufweisen.

- 5 „Funktionale Äquivalente“ sind außerdem Fusionsproteine, welche ein der oben genannten Polypeptidsequenzen oder davon abgeleitete funktionale Äquivalente und wenigstens eine weitere, davon funktionell verschiedene, heterologe Sequenz in funktioneller N- oder C-terminaler Verknüpfung (d.h. ohne gegenseitigen wesentliche funktionelle Beeinträchtigung der Fusionsproteinteile) aufweisen. Nichtlimitierende Beispiele für derartige heterologe Sequenzen sind z.B. Signalpeptide, Enzyme, Immunoglobuline, Oberflächenantigene, Rezeptoren oder Rezeptorliganden.
- 10

- Erfindungsgemäß mit umfasst „funktionale Äquivalente“ sind Homologe zu den konkret offenbarten Proteinen. Diese besitzen, beispielsweise über die gesamte Länge, wenigstens 30%, oder etwa 40%, 50 %, vorzugsweise wenigstens etwa 60 %, 65%, 70%, oder 75% ins besondere wenigsten 85 %, wie z.B. 90%, 95% oder 99%, Homologie zu einer der konkret offenbarten Sequenzen, berechnet nach dem Algorithmus von Pearson und Lipman, Proc. Natl. Acad. Sci. (USA) 85(8), 1988, 2444-2448. Der Homologiegrad spiegelt insbesondere den Grad der Identität zwischen veränderter und nicht veränderter Sequenz wider.
- 15

- 20 Homologe der erfindungsgemäßen Proteine oder Polypeptide können durch Mutagenese erzeugt werden, z.B. durch Punktmutation oder Verkürzung des Proteins. Der Begriff „Homolog“, wie er hier verwendet wird, betrifft auch eine variante Form des Proteins, die als Agonist oder Antagonist der Protein-Aktivität wirkt.

- 25 Homologe des erfindungsgemäßen Proteine können durch Screening kombinatorischer Banken von Mutanten, wie z.B. Verkürzungsmutanten, identifiziert werden. Beispielsweise kann eine variierte Bank von Protein-Varianten durch kombinatorische Mutagenese auf Nukleinsäureebene erzeugt werden, wie z.B. durch enzymatisches Ligieren eines Gemisches synthetischer Oligonukleotide. Es gibt eine Vielzahl von Verfahren, die zur Herstellung von Banken potentieller Homologer aus einer degenerierten Oligonukleotidsequenz verwendet werden können. Die chemische Synthese einer degenerierten Gensequenz kann in einem DNA-Syntheseautomaten durchgeführt werden, und das synthetische Gen kann dann in einen geeigneten Expressions-
- 30



vektor ligiert werden. Die Verwendung eines degenerierten Gensatzes ermöglicht die Bereitstellung sämtlicher Sequenzen in einem Gemisch, die den gewünschten Satz an potentiellen Proteinsequenzen codieren. Verfahren zur Synthese degenerierter Oligonukleotide sind dem Fachmann bekannt (Z.B. Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al. (1983) Nucleic Acids Res. 11:477).

Zusätzlich können Banken von Fragmenten des Protein-Codons verwendet werden, um eine variierte Population von Protein-Fragmenten zum Screening und zur anschließenden Selektion von Homologen eines erfindungsgemäßen Proteins zu erzeugen. Bei einer Ausführungsform kann eine Bank von kodierenden Sequenzfragmenten durch Behandeln eines doppelsträngigen PCR-Fragmentes einer kodierenden Sequenz mit einer Nuklease unter Bedingungen, unter denen ein Nicking nur etwa einmal pro Molekül erfolgt, Denaturieren der doppelsträngigen DNA, Renaturieren der DNA unter Bildung doppelsträngiger DNA, die Sense-/Antisense-Paare von verschiedenen genickten Produkten umfassen kann, Entfernen einzelsträngiger Abschnitte aus neu gebildeten Duplices durch Behandlung mit S1-Nuclease und Ligieren der resultierenden Fragmentbank in einen Expressionsvektor erzeugt werden. Durch dieses Verfahren kann eine Expressionsbank hergeleitet werden, die N-terminale, C-terminale und interne Fragmente mit verschiedenen Größen des erfindungsgemäßen Proteins kodiert.

Im Stand der Technik sind mehrere Techniken zum Screening von Genprodukten kombinatorischer Banken, die durch Punktmutationen oder Verkürzung hergestellt worden sind, und zum Screening von cDNA-Banken auf Genprodukte mit einer ausgewählten Eigenschaft bekannt. Diese Techniken lassen sich an das schnelle Screening der Genbanken anpassen, die durch kombinatorische Mutagenese erfindungsgemäßer Homologer erzeugt worden sind. Die am häufigsten verwendeten Techniken zum Screening großer Genbanken, die einer Analyse mit hohem Durchsatz unterliegen, umfassen das Klonieren der Genbank in replizierbare Expressionsvektoren, Transformieren der geeigneten Zellen mit der resultierenden Vektorenbank und Exprimieren der kombinatorischen Gene unter Bedingungen, unter denen der Nachweis der gewünschten Aktivität die Isolation des Vektors, der das Gen codiert, dessen Produkt nachgewiesen wurde, erleichtert. Recursive-Ensemble-Mutagenese (REM), eine Technik, die die Häufigkeit funktioneller Mutanten in den Banken vergrößert, kann in Kombination mit den Screeningtests verwendet

werden, um Homologe zu identifizieren (Arkin und Yourvan (1992) PNAS 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331

c) Erfindungsgemäße Polynukleotide

5

Gegenstand der Erfindung sind ebenso Nukleinsäuresequenzen (einzeln- und doppelsträngige DNA- und RNA-Sequenzen, wie z.B. cDNA und mRNA), kodierend für eines der obigen meth-Enzyme und deren funktionalen Äquivalenten, welche z.B. auch unter Verwendung künstlicher Nukleotidanaloga zugänglich sind.

10

Die Erfindung betrifft sowohl isolierte Nukleinsäuremoleküle, welche für erfindungsgemäße Polypeptide bzw. Proteine oder biologisch aktive Abschnitte davon kodieren, sowie Nukleinsäurefragmente, die z.B. zur Verwendung als Hybridisierungs-sonden oder Primer zur Identifizierung oder Amplifizierung von erfindungsgemäßer kodierenden Nukleinsäuren verwendet werden können.

15

Die erfindungsgemäßen Nukleinsäuremoleküle können zudem untranslatierte Sequenzen vom 3'- und/oder 5'-Ende des kodierenden Genbereichs enthalten

20

Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure zugegen sind und kann überdies im wesentlichen frei von anderem zellulären Material oder Kulturmedium sein, wenn es durch rekombinante Techniken hergestellt wird, oder frei von chemischen Vorstufen oder anderen Chemikalien sein, wenn es chemisch synthetisiert wird.

25

Die Erfindung umfasst weiterhin die zu den konkret beschriebenen Nukleotidsequenzen komplementären Nukleinsäuremoleküle oder einen Abschnitt davon.

30

Die erfindungsgemäßen Nukleotidsequenzen ermöglichen die Erzeugung von Sonden und Primern, die zur Identifizierung und/oder Klonierung von homologer Sequenzen in anderen Zelltypen und Organismen verwendbar sind. Solche Sonden bzw. Primer umfassen gewöhnlich einen Nukleotidsequenzbereich, der unter stringenten Bedingungen an mindestens etwa 12, vorzugsweise mindestens etwa 25, wie z.B. etwa 40, 50 oder 75 aufeinanderfolgende Nukleotide eines

Sense-Stranges einer erfindungsgemäßen Nukleinsäuresequenz oder eines entsprechenden Antisense-Stranges hybridisiert.

Weitere erfindungsgemäße Nukleinsäuresequenzen sind abgeleitet von SEQ ID NO:1, 3, 5, 7, 9,  
5 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49 oder 51 und unterscheiden sich davon durch Addition, Substitution, Insertion oder Deletion einzelner oder mehrerer Nukleotide, kodieren aber weiterhin für Polypeptide mit dem gewünschten Eigenschaftsprofil. Dies können Polynukleotide sein, die zu obigen Sequenzen, beispielsweise über die gesamte Länge, in mindestens etwa 50%, 55%, 60%, 65%, 70%, 80% oder 90%, vorzugsweise in mindestens etwa 95%, 96%, 97%, 98% oder 99% der Sequenzpositionen identisch sind.  
10

Erfindungsgemäß umfasst sind auch solche Nukleinsäuresequenzen, die sogenannte stumme Mutationen umfassen oder entsprechend der Codon-Nutzung eines speziellen Ursprungs- oder Wirtsorganismus, im Vergleich zu einer konkret genannten Sequenz verändert sind, ebenso wie  
15 natürlich vorkommende Varianten, wie z.B. Spleißvarianten oder Allelvarianten, davon. Gegenstand sind ebenso durch konservative Nukleotidsubstitutionen (d.h. die betreffende Aminosäure wird durch eine Aminosäure gleicher Ladung, Größe, Polarität und/oder Löslichkeit ersetzt) erhältliche Sequenzen.

20 Gegenstand der Erfindung sind auch die durch Sequenzpolymorphismen von den konkret offenbarten Nukleinsäuren abgeleiteten Moleküle. Diese genetischen Polymorphismen können zwischen Individuen innerhalb einer Population aufgrund der natürlichen Variation existieren. Diese natürlichen Variationen bewirken üblicherweise eine Varianz von 1 bis 5 % in der Nukleotidsequenz eines Gens.

25 Weiterhin umfasst die Erfindung auch Nukleinsäuresequenzen, welchen mit oben genannten kodierenden Sequenzen hybridisieren oder dazu komplementär sind. Diese Polynukleotide lassen sich bei Durchmusterung von genomischen oder cDNA-Banken auffinden und gegebenenfalls daraus mit geeigneten Primern mittels PCR vermehren und anschließend beispielsweise mit geeigneten Sonden isolieren. Eine weitere Möglichkeit bietet die Transformation geeigneter  
30 Mikroorganismen mit erfindungsgemäßen Polynukleotiden oder Vektoren, die Vermehrung der Mikroorganismen und damit der Polynukleotide und deren anschließende Isolierung. Darüber hinaus können erfindungsgemäße Polynukleotide auch auf chemischem Wege synthetisiert

werden.

- Unter der Eigenschaft, an Polynukleotide „hybridisieren“ zu können, versteht man die Fähigkeit eines Poly- oder Oligonukleotids unter stringenten Bedingungen an eine nahezu komplementäre Sequenz zu binden, während unter diesen Bedingungen unspezifische Bindungen zwischen nicht-komplementären Partnern unterbleiben. Dazu sollten die Sequenzen zu 70-100%, vorzugsweise zu 90-100%, komplementär sein. Die Eigenschaft komplementärer Sequenzen, spezifisch aneinander binden zu können, macht man sich beispielsweise in der Northern- oder Southern-Blot-Technik oder bei der Primerbindung in PCR oder RT-PCR zunutze. Üblicherweise werden dazu Oligonukleotide ab einer Länge von 30 Basenpaaren eingesetzt. Unter stringenten Bedingungen versteht man beispielsweise in der Northern-Blot-Technik die Verwendung einer 50 – 70 °C, vorzugsweise 60 – 65 °C warmen Waschlösung, beispielsweise 0,1x SSC-Puffer mit 0,1% SDS (20x SSC: 3M NaCl, 0,3M Na-Citrat, pH 7,0) zur Elution unspezifisch hybridisierter cDNA-Sonden oder Oligonukleotide. Dabei bleiben, wie oben erwähnt, nur in hohem Maße komplementäre Nukleinsäuren aneinander gebunden. Die Einstellung stringenter Bedingungen ist dem Fachmann bekannt und ist z.B. in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. beschrieben.

c) Isolierung der kodierenden methH-Gene

20

Die für das Enzym Methionin-Synthase (EC 2.1.1.13) kodierenden methH Gene aus den Organismen obiger Liste I sind in an sich bekannter Weise isolierbar.

- Zur Isolierung der methH-Gene oder auch anderer Gene der Organismen aus obiger Liste I wird zunächst eine Genbank dieses Organismus in Escherichia coli (E. coli) angelegt. Das Anlegen von Genbanken ist in allgemein bekannten Lehrbüchern und Handbüchern ausführlich beschrieben. Als Beispiel seien das Lehrbuch von Winnacker: Gene und Klone, Eine Einführung in die Gentechnologie (Verlag Chemie, Weinheim, Deutschland, 1990), oder das Handbuch von Sambrook et al.: Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1989) genannt. Eine sehr bekannte Genbank ist die des E. coli K-12 Stammes W3110, die von Kohara et al. (Cell 50, 495-508 (198)) in  $\lambda$ -Vektoren angelegt wurde.

Zur Herstellung einer Genbank von Organismen der Liste I in *E. coli* können Cosmide, wie der Cosmidvektor SuperCos I (Wahl et al., 1987, Proceedings of the National Academy of Sciences USA, 84: 2160-2164), aber auch Plasmide, wie pBR322 (BoliVal; Life Sciences, 25, 807-818 (1979)) oder pUC9 (Vieira et al., 1982, Gene, 19: 259-268), verwendet werden. Als Wirte eignen sich besonders solche *E. coli* Stämme, die restriktions- und rekombinationsdefekt sind. Ein Beispiel hierfür ist der Stamm DH5 $\alpha$ mc $r$ , der von Grant et al. (Proceedings of the National Academy of Sciences USA, 87 (1990) 4645-4649) beschrieben wurde. Die mit Hilfe von Cosmiden klonierten langen DNA-Fragmente können anschließend wiederum in gängige, für die Sequenzierung geeignete Vektoren subkloniert und anschließend sequenziert werden, so wie es z. B. bei Sanger et al. (proceedings of the National Academy of Sciences of the United States of America, 74: 5463-5467, 1977) beschrieben ist.

Die erhaltenen DNA-Sequenzen können dann mit bekannten Algorithmen bzw. Sequenzanalyse-Programmen, wie z. B. dem von Staden (Nucleic Acids Research 14,217-232(1986)), dem von Marck (Nucleic Acids Research 16, 1829-1836 (1988)) oder dem GCG-Programm von Butler (Methods of Biochemical Analysis 39, 74-97 (1998)), untersucht werden.

Die für die methH-Gene kodierenden DNA-Sequenzen von Organismen gemäß obiger Liste I wurden gefunden. Insbesondere wurden DNA-Sequenzen gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49 und 51 gefunden. Weiterhin wurde aus diesen vorliegenden DNA-Sequenzen mit den oben beschriebenen Methoden die Aminosäuresequenzen der entsprechenden Proteine abgeleitet. Durch SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 und 52 sind die sich ergebenden Aminosäuresequenzen der methH Genprodukte dargestellt.

Kodierende DNA-Sequenzen, die sich aus den Sequenzen gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49 und 51 durch die Degeneration des genetischen Kodes ergeben, sind ebenfalls Gegenstand der Erfindung. In gleicher Weise sind DNA-Sequenzen, die mit diesen Sequenzen oder davon abgeleiteten Sequenzteilen hybridisieren, Gegenstand der Erfindung.

Anleitungen zur Identifizierung von DNA-Sequenzen mittels Hybridisierung findet der Fachmann unter anderem im Handbuch "The DIG System Users Guide für Filter Hybridization" der Firma

Boehringer Mannheim GmbH (Mannheim, Deutschland, 1993) und bei Liebl et al. (International Journal of Systematic Bacteriology (1991) 41: 255-260). Anleitungen zur Amplifikation von DNA-Sequenzen mit Hilfe der Polymerase-Kettenreaktion (PCR) findet der Fachmann unter anderem im Handbuch von Gait: Oligonukleotide synthesis: A Practical Approach (IRL Press, Oxford, UK, 1984) und bei Newton und Graham: PCR (Spektrum Akademischer Verlag, Heidelberg, Deutschland, 1994).

Weiterhin ist bekannt, dass Änderungen am N- und/oder C- Terminus eines Proteins dessen Funktion nicht wesentlich beeinträchtigen oder sogar stabilisieren können. Angaben hierzu findet der Fachmann unter anderem bei Ben-Bassat et al. (Journal of Bacteriology 169: 751-757 (1987)), bei O'Regan et al. (Gene 77: 237-251 (1989)), bei Sahin-Toth et al. (Protein Sciences 3: 240-247 (1994)), bei Hochuli et al. (Biotechnology 6: 1321-1325 (1988)) und in bekannten Lehrbüchern der Genetik und Molekularbiologie.

Aminosäuresequenzen, die sich in entsprechender Weise aus den SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 und 52 ergeben, sind ebenfalls Bestandteil der Erfindung.

d) Erfindungsgemäß verwendete Wirtszellen

Weitere Gegenstände der Erfindung betreffen als Wirtszelle dienende Mikroorganismen, insbesondere coryneforme Bakterien, die einen Vektor, insbesondere Pendelvektor oder Plasmidvektor, der wenigstens ein meth Gen erfindungsgemäßer Definition trägt, enthalten oder in denen ein erfindungsgemäßes meth Gen exprimiert bzw. verstärkt ist.

Diese Mikroorganismen können schwefelhaltige Feinchemikalien, insbesondere L-Methionin, aus Glucose, Saccharose, Lactose, Fructose, Maltose, Melasse, Stärke, Cellulose oder aus Glycerin und Ethanol herstellen. Vorzugsweise sind dies coryneforme Bakterien, insbesondere der Gattung Corynebacterium. Aus der Gattung Corynebacterium ist insbesondere die Art Corynebacterium glutamicum zu nennen, die in der Fachwelt für ihre Fähigkeit bekannt ist, L-Aminosäuren zu produzieren.

Als Beispiele für geeignete Stämme coryneformer Bakterien sind solche der Gattung *Corynebacterium*, insbesondere der Art *Corynebacterium glutamicum* (*C. glutamicum*), wie

*Corynebacterium glutamicum* ATCC 13032,

*Corynebacterium acetoglutamicum* ATCC 15806,

5 *Corynebacterium acetoacidophilum* ATCC 13870,

*Corynebacterium thermoaminogenes* FERM BP-1539,

*Corynebacterium melassecola* ATCC 17965

oder

10 der Gattung *Brevibacterium*, wie

*Brevibacterium flavum* ATCC 14067

*Brevibacterium lactofermentum* ATCC 13869 und

*Brevibacterium divaricatum* ATCC 14020 zu nennen;

oder davon abgeleitete Stämme, wie

15 *Corynebacterium glutamicum* KFCC10065

*Corynebacterium glutamicum* ATCC21608

welche ebenfalls die gewünschte Feinchemikalie oder deren Vorstufe(n) produzieren.

Mit der Abkürzung KFCC ist die Korean Federation of Culture Collection gemeint, mit der Abkür-

20 zung ATCC die American type strain culture collection und mit der Abkürzung FERM die Sammlung des National institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Japan.

e) Durchführung der erfindungsgemäßen Fermentation

25

Erfindungsgemäß wurde festgestellt, dass coryneforme Bakterien nach Überexpression eines *methH* Gens aus Organismen der Liste I in vorteilhafter Weise schwefelhaltige Feinchemikalien, insbesondere L-Methionin, produzieren.

30 Zur Erzielung einer Überexpression kann der Fachmann unterschiedliche Maßnahmen einzeln oder in Kombination ergreifen. So kann die Kopienzahl der entsprechenden Gene erhöht werden, oder es kann die Promotor- und Regulationsregion oder die Ribosomenbindungsstelle, die sich stromaufwärts des Strukturgens befindet, mutiert werden. In gleicher Weise wirken Expres-

- sionskassetten, die stromaufwärts des Strukturgens eingebaut werden. Durch induzierbare Promotoren ist es zusätzlich möglich, die Expression im Verlaufe der fermentativen L-Methionin-Produktion zu steigern. Durch Maßnahmen zur Verlängerung der Lebensdauer der mRNA wird ebenfalls die Expression verbessert. Weiterhin wird durch Verhinderung des Abbaus des Enzymproteins ebenfalls die Enzymaktivität verstärkt. Die Gene oder Genkonstrukte können entweder in Plasmiden mit unterschiedlicher Kopienzahl vorliegen oder im Chromosom integriert und amplifiziert sein. Alternativ kann weiterhin eine Überexpression der betreffenden Gene durch Veränderung der Medienzusammensetzung und Kulturführung erreicht werden.
- 5
- 10 Anleitungen hierzu findet der Fachmann unter anderem bei Martin et al. (Bionotechnology 5, 137-146 (1987)), bei Guerrero et al. (Gene 138, 35-41 (1994)), Tsuchiya und Morinaga (Bio/Technology 6, 428-430 (1988)), bei Eikmanns et al. (Gene 102, 93-98 (1991)), in der Europäischen Patentschrift 0472869, im US Patent 4,601,893, bei Schwarzer und Pühler (Biotechnology 9, 84-87 (1991)), bei Remscheid et al. (Applied and Environmental Microbiology 60, 126-132
- 15 (1994)), bei LaBarre et al. (Journal of Bacteriology 175, 1001-1007 (1993)), in der Patentanmeldung WO 96/15246, bei Malumbres et al. (Gene 134, 15-24 (1993)), in der japanischen Offenlegungsschrift JP-A-10-229891, bei Jensen und Hammer (Biotechnology and Bioengineering 58, 191-195 (1998)), bei Makrides (Microbiological Reviews 60 : 512-538 (1996)) und in bekannten Lehrbüchern der Genetik und Molekularbiologie.
- 20
- Gegenstand der Erfindung sind deshalb auch Expressionskonstrukte, enthaltend unter der genetischen Kontrolle regulativer Nukleinsäuresequenzen eine für ein erfindungsgemäßes Polypeptid kodierende Nukleinsäuresequenz; sowie Vektoren, umfassend wenigstens eines dieser Expressionskonstrukte. Vorzugsweise umfassen solche erfindungsgemäßen Konstrukte 5'-stromaufwärts von der jeweiligen kodierenden Sequenz einen Promotor und 3'-stromabwärts eine Terminatorsequenz sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils operativ verknüpft mit der kodierenden Sequenz. Unter einer „operativen Verknüpfung“ versteht man die sequentielle Anordnung von Promotor, kodierender Sequenz, Terminator und gegebenenfalls weiterer regulativer Elemente derart, dass jedes der regulativen Elemente seine
- 25
- 30 Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann. Beispiele für operativ verknüpfbare Sequenzen sind Aktivierungssequenzen sowie Enhancer und dergleichen. Weitere regulative Elemente umfassen selektierbare Marker, Amplifikationssignale, Replikationsursprünge und dergleichen. Geeignete regulatorische Sequenzen sind z.B. be-



schrieben in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990).

5 Zusätzlich zu den artifiziellen Regulationssequenzen kann die natürliche Regulationssequenz vor dem eigentlichen Strukturgen noch vorhanden sein. Durch genetische Veränderung kann diese natürliche Regulation gegebenenfalls ausgeschaltet und die Expression der Gene erhöht oder erniedrigt werden. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das heißt es werden keine zusätzlichen Regulationssignale vor das Strukturgen insertiert und der natürliche Promotor mit seiner Regulation wird nicht entfernt. Statt dessen wird die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert oder verringert wird. Die Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

15 Beispiele für brauchbare Promotoren sind: die Promotoren, *ddh*, *amy*, *lysC*, *dapA*, *lysA* aus *Corynebacterium glutamicum*, aber auch gram-positiven Promotoren SPO2 wie sie in *Bacillus Subtilis* and Its Closest Relatives, Sonenshein, Abraham L., Hoch, James A., Losick, Richard; ASM Press, District of Columbia, Washington und Patek M. Eikmanns B.J. Patek J. Sahm H. Microbiology. 142 1297-309, 1996 beschrieben sind, oder aber auch *cos*-, *tac*-, *trp*-, *tet*-, *trp-tet*-, *lpp*-, *lac*-, *lpp-lac*-, *lacIq*-, T7-, T5-, T3-, *gal*-, *trc*-, *ara*-, SP6-, *lambda-PR*- oder *lambda-PL*-Promotor, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden. Bevorzugt ist auch die Verwendung induzierbarer Promotoren, wie z.B. licht- und insbesondere temperaturinduzierbarer Promotoren, wie der *P<sub>P</sub>*-Promotor. Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen verwendet werden. Darüber hinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

25 Die genannten regulatorischen Sequenzen sollen die gezielte Expression der Nukleinsäuresequenzen ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird.

30 Die regulatorischen Sequenzen bzw. Faktoren können dabei vorzugsweise die Expression positiv beeinflussen und dadurch erhöhen oder erniedrigen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke

Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

- 5 Die Herstellung einer Expressionskassette erfolgt durch Fusion eines geeigneten Promotors, einer geeigneten Shine-Dalgarno-Sequenz mit einer methH-Nukleotidsequenz sowie einem geeigneten Terminationssignal. Dazu verwendet man gängige Rekombinations- und Klonierungstechniken, wie sie beispielsweise in Current Protocols in Molecular Biology, 1993, John Wiley & Sons, Incorporated, New York New York, PCR Methods, Gelfand, David H., Innis, Michael A., Sninsky, John J. 1999, Academic Press, Incorporated, California, San Diego, .. PCR Cloning Protocols, Methods in Molecular Biology Ser., Vol. 192, 2nd ed., Humana Press, New Jersey, Totowa. T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987) beschrieben sind.

- Das rekombinante Nukleinsäurekonstrukt bzw. Genkonstrukt wird zur Expression in einem geeigneten Wirtsorganismus vorteilhafterweise in einen wirtsspezifischen Vektor insertiert, der eine optimale Expression der Gene im Wirt ermöglicht. Vektoren sind dem Fachmann wohl bekannt und können beispielsweise aus "Cloning Vectors" (Pouwels P. H. et al., Hrsg, Elsevier, Amsterdam-New York-Oxford, 1985) entnommen werden. Unter Vektoren sind außer Plasmiden auch alle anderen dem Fachmann bekannten Vektoren, wie beispielsweise Phagen, Transposons, IS-Elemente, Phasmide, Cosmide, und lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können autonom im Wirtsorganismus repliziert oder chromosomal repliziert werden.

- Zur Verstärkung wurden erfindungsgemäße methH Gene beispielhaft mit Hilfe von episomalen Plasmiden überexprimiert. Als Plasmide eignen sich solche, die in coryneformen Bakterien repliziert werden. Zahlreiche bekannte Plasmidvektoren, wie z. B. pZ1 (Menkel et al., Applied and Environmental Microbiology (1989) 64: 549-554), pEKEx1 (Eikmanns et al., Gene 102: 93-98 (1991)) oder pHS2-1 (Sonnen et al., Gene 107: 69-74 (1991)) beruhen auf den kryptischen Plasmiden pHM1519, pBL1 oder pGA1. Andere Plasmidvektoren, wie z. B. pCLiK5MCS, oder solche, die auf pCG4 (US-A 4,489,160) oder pNG2 (Serwold-Davis et al., FEMS Microbiology

Letters 66, 119-124 (1990)) oder pAG1 (US-A 5,158,891) beruhen, können in gleicher Weise verwendet werden.

- 5 Weiterhin eignen sich auch solche Plasmidvektoren mit Hilfe derer man das Verfahren der Genamplifikation durch Integration in das Chromosom anwenden kann, so wie es beispielsweise von Remscheid et al. (Applied and Environmental Microbiology 60,126-132 (1994)) zur Duplikation bzw. Amplifikation des *hom-thrB*-Operons beschrieben wurde. Bei dieser Methode wird das vollständige Gen in einen Plasmidvektor kloniert, der in einem Wirt (typischerweise *E. coli*), nicht aber in *C. glutamicum* replizieren kann. Als Vektoren kommen beispielsweise pSUP301 (Simon
- 10 et al., Bio/ Technology 1,784-791 (1983)), pK18mob oder pK19mob (Schäfer et al., Gene 145,69-73 (1994)), Bernard et al., Journal of Molecular Biology, 234: 534-541 (1993)), pEM1 (Schrumpf et al. 1991, Journal of Bacteriology 173: 4510-4516) oder pBGS8 (Spratt et al., 1986, Gene 41: 337-342) in Frage. Der Plasmidvektor, der das zu amplifizierende Gen enthält, wird anschließend durch Transformation in den gewünschten Stamm von *C. glutamicum* überführt.
- 15 Methoden zur Transformation sind beispielsweise bei Thierbach et al. (Applied Microbiology and Biotechnology 29, 356-362 (1988)), Dunican und Shivnan (Biotechnology 7, 1067-1070 (1989)) und Tauch et al. (FEMS Microbiological Letters 123,343-347 (1994)) beschrieben.

- Enzyme können durch Mutationen in den korrespondierenden Genen derart in ihrer Aktivität
- 20 beeinflusst werden, dass es zu einer teilweisen oder vollständigen Verringerung der Reaktionsgeschwindigkeit der enzymatischen Reaktion kommt. Beispiele für solche Mutationen sind dem Fachmann bekannt (Motoyama H. Yano H. Terasaki Y. Anazawa H. Applied & Environmental Microbiology. 67:3064-70, 2001, Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94.)

- 25 Zusätzlich kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben einer Expression bzw. Verstärkung eines erfindungsgemäßen *metH*-Gen eines oder mehrere Enzyme des Methionin-Biosyntheseweges oder eines damit assoziierten (d.h. in einem funktionelle Zusammenhang stehenden) Biosynthese- oder sonstigen
- 30 Stoffwechselweges, wie des Cystein-, Lysin- oder Threonin-Stoffwechselwegs, wie insbesondere der Aspartatsemialdehyd-Synthese, der Glykolyse, der Anaplerotik, des Pentose-Phosphat-Stoffwechsels, des Zitronensäure-Zyklus oder des Aminosäure-Exports zu verstärken.

## 23

So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, eines oder mehrere der folgenden Gene verstärkt sein, (d.h. z.B. in einer höheren Kopienzahl vorliegen oder ein Enzym mit höherer Aktivität oder Spezifität kodieren):

- das für eine Aspartatkinase kodierende Gen *lysC* (EP 1 108 790 A2; DNA-SEQ NO. 281),
- 5 - das für eine Aspartat-Semialdehyd Dehydrogenase kodierende Gen *asd* (EP 1 108 790 A2; DNA-SEQ NO. 282),
- das für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierende Gen *gap* (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die 3-Phosphoglycerat Kinase kodierende Gen *pgk* (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- 10 - das für die Pyruvat Carboxylase kodierende Gen *pyc* (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die Triosephosphat Isomerase kodierende Gen *tpi* (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- 15 - das für die Homoserin O-Acetyltransferase kodierende Gen *metA* (EP 1 108 790 A2; DNA-SEQ NO. 725),
- das für die Cystathionin-gamma-Synthase kodierende Gen *metB* (EP 1 108 790 A2; DNA-SEQ NO. 3491),
- das für die Cystathionin-gamma-Lyase kodierende Gen *metC* (EP 1 108 790 A2; DNA-SEQ NO. 3061),
- 20 - das für die Serin-Hydroxymethyltransferase kodierende Gen *glyA* (EP 1 108 790 A2; DNA-SEQ NO. 1110),
- das für die O-Acetylhomoserin-Sulphydrylase kodierende Gen *metY* (EP 1 108 790 A2; DNA-SEQ NO. 726),
- 25 - das für die Methylentetrahydrofolat-Reduktase kodierende Gen *metF* (EP 1 108 790 A2; DNA-SEQ NO. 2379),
- das für die Phosphoserin-Aminotransferase kodierende Gen *serC* (EP 1 108 790 A2; DNA-SEQ NO. 928)
- eines für die Phosphoserin-Phosphatase kodierende Gen *serB* (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)
- 30 - das für die Serine Acetyl-Transferase kodierende Gen *cysE* (EP 1 108 790 A2; DNA-SEQ NO. 2818)

- das für eine Homoserin-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ NO. 1306)

5 So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, in coryneformen Bakterien, vorteilhaft sein, gleichzeitig wenigstens eines der nachfolgenden Gene zu mutieren, insbesondere so, dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch einen Stoffwechselmetaboliten in ihrer Aktivität beeinflusst werden:

- 10 - das für eine Aspartatkinase kodierende Gen lysC (EP 1 108 790 A2; DNA-SEQ NO. 281),  
- das für die Pyruvat Carboxylase kodierende Gen pyc (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),  
- das für die Homoserin O-Acetyltransferase kodierende Gen metA (EP 1 108 790 A2; DNA-SEQ NO. 725),
- 15 - das für die Cystathionin-gamma-Synthase kodierende Gen metB (EP 1 108 790 A2; DNA-SEQ NO. 3491),  
- das für die Cystathionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),  
- das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ NO. 1110),
- 20 - das für die O-Acetylhomoserin-Sulphydrylase kodierende Gen metY (EP 1 108 790 A2; DNA-SEQ NO. 726),  
- das für die Methylentetrahydrofolat-Reduktase kodierende Gen metF (EP 1 108 790 A2; DNA-SEQ NO. 2379),
- 25 - das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-SEQ NO. 928)  
- eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)  
- das für die Serine Acetyl-Transferase kodierende Gen cysE (EP 1 108 790 A2; DNA-SEQ NO. 2818)
- 30 - das für eine Homoserin-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ NO. 1306)

## 25

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsgemäßen methH-Gene eines oder mehrere der folgenden Gene abzuschwächen, insbesondere deren Expression zu verringern, oder auszuschalten:

5

- das für die Homoserine-Kinase kodierende Gen thrB (EP 1 108 790 A2; DNA-SEQ NO. 3453)
- das für die Threonin Dehydratase kodierende Gen ilvA (EP 1 108 790 A2; DNA-SEQ NO. 2328)

- das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)

10

- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2; DNA-SEQ NO. 3494)

- das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)

- das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)

15

- das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)

- das für die Dihydrodipicolinat Synthase kodierende Gen dapA (EP 1 108 790 A2; DNA-SEQ NO. 3476)

- das für die Dihydrodipicolinat Reduktase kodierende Gen dapB (EP 1 108 790 A2; DNA-SEQ NO. 3477)

20

- das für die Diaminopicolinat Decarboxylase kodierende Gen lysA (EP 1 108 790 A2; DNA-SEQ NO. 3451)

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsgemäßen methH-Gene in Coryneformen Bakterien gleichzeitig wenigstens eines der folgenden Gene so zu mutieren, dass die enzymatische Aktivität des korrespondierenden Proteins teilweise oder vollständig verringert wird:

25

- das für die Homoserine-Kinase kodierende Gen thrB (EP 1 108 790 A2; DNA-SEQ NO. 3453)
- das für die Threonin Dehydratase kodierende Gen ilvA (EP 1 108 790 A2; DNA-SEQ NO. 2328)

- das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)

30

- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2; DNA-SEQ NO. 3494)
- das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)
- 5 - das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)
- das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)
- das für die Dihydrodipicolinat Synthase kodierende Gen dapA (EP 1 108 790 A2; DNA-SEQ NO. 3476)
- 10 - das für die Dihydrodipicolinat Reduktase kodierende Gen dapB (EP 1 108 790 A2; DNA-SEQ NO. 3477)
- das für die Diaminopicolinat Decarboxylase kodierende Gen lysA (EP 1 108 790 A2; DNA-SEQ NO. 3451)
- 15 Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben der Expression bzw. Verstärkung eines erfindungsgemäßen methH-Gens unerwünschte Nebenreaktionen auszuschalten, welche beispielsweise die Ausbeute an der Feinchemikalie verringern (Nakayama: "Breeding of Amino Acid Producing Microorganisms", In: Overproduction of Microbial Products, Krumphanz, Sikyta, Vanek (eds.), Academic Press, London, UK, 1982).
- 20

- Die erfindungsgemäß hergestellten Mikroorganismen können kontinuierlich oder diskontinuierlich im batch-Verfahren (Satzkultivierung) oder im fed batch (Zulaufverfahren) oder repeated fed batch Verfahren (repetitives Zulaufverfahren) zur Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, kultiviert werden. Eine Zusammenfassung über bekannte Kultivierungsmethoden ist im Lehrbuch von Chmiel (Bioprozeßtechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) oder im Lehrbuch von Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)) zu finden.
- 25
- 30 Das zu verwendende Kulturmedium hat in geeigneter Weise den Ansprüchen der jeweiligen Stämme zu genügen. Beschreibungen von Kulturmedien verschiedener Mikroorganismen sind im Handbuch "Manual of Methods für General Bacteriology" der American Society für Bacteriology (Washington D. C., USA, 1981) enthalten.

Diese erfindungsgemäß einsetzbaren Medien umfassen gewöhnlich eine oder mehrere Kohlenstoffquellen, Stickstoffquellen, anorganische Salze, Vitamine und/oder Spurenelemente.

- 5 Bevorzugte Kohlenstoffquellen sind Zucker, wie Mono-, Di- oder Polysaccharide. Sehr gute Kohlenstoffquellen sind beispielsweise Glucose, Fructose, Mannose, Galactose, Ribose, Sorbose, Ribulose, Lactose, Maltose, Saccharose, Raffinose, Stärke oder Cellulose. Man kann Zucker auch über komplexe Verbindungen, wie Melassen, oder andere Nebenprodukte der Zucker-Raffinierung zu den Medien geben. Es kann auch vorteilhaft sein, Gemische verschiedener Kohlenstoffquellen zuzugeben. Andere mögliche Kohlenstoffquellen sind Öle und Fette wie z. B. Sojaöl, Sonnenblumenöl, Erdnußöl und Kokosfett, Fettsäuren wie z. B. Palmitinsäure, Stearinsäure oder Linolsäure, Alkohole wie z. B. Glycerin, Methanol oder Ethanol und organische Säuren wie z. B. Essigsäure oder Milchsäure.
- 10
- 15 Stickstoffquellen sind gewöhnlich organische oder anorganische Stickstoffverbindungen oder Materialien, die diese Verbindungen enthalten. Beispielhafte Stickstoffquellen umfassen Ammoniak-Gas oder Ammoniumsalze, wie Ammoniumsulfat, Ammoniumchlorid, Ammoniumphosphat, Ammoniumcarbonat oder Ammoniumnitrat, Nitrate, Harnstoff, Aminosäuren oder komplexe Stickstoffquellen, wie Maisquellwasser, Sojamehl, Sojaprotein, Hefeextrakt, Fleischextrakt und
- 20 andere. Die Stickstoffquellen können einzeln oder als Mischung verwendet werden.

Anorganische Salzverbindungen, die in den Medien enthalten sein können, umfassen die Chlorid-, Phosphor- oder Sulfatsalze von Calcium, Magnesium, Natrium, Kobalt, Molybdän, Kalium, Mangan, Zink, Kupfer und Eisen

25

Als Schwefelquelle für die Herstellung von schwefelhaltigen Feinchemikalien, insbesondere von Methionin, können anorganische schwefelhaltige Verbindungen wie beispielsweise Sulfate, Sulfite, Dithionite, Tetrathionate, Thiosulfate, Sulfide aber auch organische Schwefelverbindungen, wie Mercaptane und Thiole, verwendet werden.

30

Als Phosphorquelle können Phosphorsäure, Kaliumdihydrogenphosphat oder Dikaliumhydrogenphosphat oder die entsprechenden Natrium haltigen Salze verwendet werden.



Chelatbildner können zum Medium gegeben werden, um die Metallionen in Lösung zu halten. Besonders geeignete Chelatbildner umfassen Dihydroxyphenole, wie Catechol oder Protocatechuat, oder organische Säuren, wie Citronensäure.

- 5 Die erfindungsgemäß eingesetzten Fermentationsmedien enthalten üblicherweise auch andere Wachstumsfaktoren, wie Vitamine oder Wachstumsförderer, zu denen beispielsweise Biotin, Riboflavin, Thiamin, Folsäure, Nikotinsäure, Panthothenat und Pyridoxin gehören. Wachstumsfaktoren und Salze stammen häufig von komplexen Medienkomponenten, wie Hefeextrakt, Melassen, Maisquellwasser und dergleichen. Dem Kulturmedium können überdies geeignete Vor-
- 10 stufen zugesetzt werden. Die genaue Zusammensetzung der Medienverbindungen hängt stark vom jeweiligen Experiment ab und wird für jeden spezifischen Fall individuell entschieden. Information über die Medienoptimierung ist erhältlich aus dem Lehrbuch "Applied Microbiol. Physiology, A Practical Approach" (Hrsg. P.M. Rhodes, P.F. Stanbury, IRL Press (1997) S. 53-73, ISBN 0 19 963577 3). Wachstumsmedien lassen sich auch von kommerziellen Anbietern beziehen, wie Standard 1 (Merck) oder BHI (Brain heart infusion, DIFCO) und dergleichen.
- 15

- Sämtliche Medienkomponenten werden, entweder durch Hitze (20 min bei 1,5 bar und 121°C) oder durch Sterilfiltration, sterilisiert. Die Komponenten können entweder zusammen oder nötigenfalls getrennt sterilisiert werden. Sämtliche Medienkomponenten können zu Beginn der An-
- 20 zucht zugegen sein oder wahlfrei kontinuierlich oder chargenweise hinzugegeben werden.

- Die Temperatur der Kultur liegt normalerweise zwischen 15°C und 45°C, vorzugsweise bei 25°C bis 40°C und kann während des Experimentes konstant gehalten oder verändert werden. Der pH-Wert des Mediums sollte im Bereich von 5 bis 8,5, vorzugsweise um 7,0 liegen. Der pH-Wert
- 25 für die Anzucht läßt sich während der Anzucht durch Zugabe von basische Verbindungen wie Natriumhydroxid, Kaliumhydroxid, Ammoniak bzw. Ammoniakwasser oder saure Verbindungen wie Phosphorsäure oder Schwefelsäure kontrollieren. Zur Kontrolle der Schaumentwicklung können Antischaummittel, wie z. B. Fettsäurepolyglykolester, eingesetzt werden. Zur Aufrechterhaltung der Stabilität von Plasmiden können dem Medium geeignete selektiv wirkende Stoffe,
- 30 wie z. B. Antibiotika, hinzugefügt werden. Um aerobe Bedingungen aufrechtzuerhalten, werden Sauerstoff oder Sauerstoff haltige Gasmischungen, wie z. B. Umgebungsluft, in die Kultur eingetragen. Die Temperatur der Kultur liegt normalerweise bei 20°C bis 45°C. Die Kultur wird solange fortgesetzt, bis sich ein Maximum des gewünschten Produktes gebildet hat. Dieses Ziel wird

normalerweise innerhalb von 10 Stunden bis 160 Stunden erreicht.

Die so erhaltenen, insbesondere L-Methionin enthaltenden, Fermentationsbrühen haben üblicherweise eine Trockenmasse von 7,5 bis 25 Gew.-%.

5

Vorteilhaft ist außerdem auch, wenn die Fermentation zumindest am Ende, insbesondere jedoch über mindestens 30% der Fermentationsdauer zuckerlimitiert gefahren wird. Das heißt, dass während dieser Zeit die Konzentration an verwertbarem Zucker im Fermentationsmedium auf  $\geq 0$  bis 3 g/l gehalten, beziehungsweise abgesenkt wird.

10

Die Fermentationsbrühe wird anschließend weiterverarbeitet. Je nach Anforderung kann die Biomasse ganz oder teilweise durch Separationsmethoden, wie z. B. Zentrifugation, Filtration, Dekantieren oder einer Kombination dieser Methoden aus der Fermentationsbrühe entfernt oder vollständig in ihr belassen werden.

15

Anschließend kann die Fermentationsbrühe mit bekannten Methoden, wie z. B. mit Hilfe eines Rotationsverdampfers, Dünnschichtverdampfers, Fallfilmverdampfers, durch Umkehrosmose, oder durch Nanofiltration, eingedickt beziehungsweise aufkonzentriert werden. Diese aufkonzentrierte Fermentationsbrühe kann anschließend durch Gefriertrocknung, Sprühtrocknung, 20 Sprühgranulation oder durch anderweitige Verfahren aufgearbeitet werden.

Es ist aber auch möglich die schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, weiter aufzureinigen. Hierzu wird die produktthaltige Brühe nach dem Abtrennen der Biomasse einer Chromatographie mit einem geeigneten Harz unterworfen, wobei das gewünschte Produkt oder 25 die Verunreinigungen ganz oder teilweise auf dem Chromatographieharz zurückgehalten werden. Diese Chromatographieschritte können nötigenfalls wiederholt werden, wobei die gleichen oder andere Chromatographieharze verwendet werden. Der Fachmann ist in der Auswahl der geeigneten Chromatographieharze und ihrer wirksamsten Anwendung bewandert. Das gereinigte Produkt kann durch Filtration oder Ultrafiltration konzentriert und bei einer Temperatur aufbe- 30 wahrt werden, bei der die Stabilität des Produktes maximal ist.

Die Identität und Reinheit der isolierten Verbindung(en) kann durch Techniken des Standes der Technik bestimmt werden. Diese umfassen Hochleistungs-Flüssigkeitschromatographie (HPLC),

- spektroskopische Verfahren, Färbeverfahren, Dünnschichtchromatographie, NIRS, Enzymtest oder mikrobiologische Tests. Diese Analyseverfahren sind zusammengefaßt in: Patek et al. (1994) Appl. Environ. Microbiol. 60:133-140; Malakhova et al. (1996) Biotekhnologiya 11 27-32; und Schmidt et al. (1998) Bioprocess Engineer. 19:67-70. Ulmann's Encyclopedia of Industrial Chemistry (1996) Bd. A27, VCH: Weinheim, S. 89-90, S. 521-540, S. 540-547, S. 559-566, 575-581 und S. 581-587; Michal, G (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. et al. (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17.

- 10 Die Erfindung wird nun anhand der folgenden nicht-limitierenden Beispiele näher beschrieben:

**Beispiel 1: Konstruktion von pCLIK5MCS**

- 15 Zunächst wurden Ampicillinresistenz und Replikationsursprung des Vektors pBR322 mit den Oligonukleotiden p1.3 (SEQ ID NO:53) und p2.3 (SEQ ID NO:54) mit Hilfe der Polymerase-Kettenreaktion (PCR) amplifiziert.

p1.3 (SEQ ID NO:53)

5'-CCCGGGATCCGCTAGCGGCGCGCCGGCCGGCCCGGTGTGAAATACCGCACAG-3'

20

p2.3 (SEQ ID NO:54)

5'-TCTAGACTCGAGCGGCCGCGGCCGGCCTTTAAATTGAAGACGAAAGGGCCTCG-3'

- 25 Neben den zu pBR322 komplementären Sequenzen, enthält das Oligonukleotid p1.3 (SEQ ID NO:53) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen SmaI, BamHI, NheI und Ascl und das Oligonukleotid p2.3 (SEQ ID NO:54) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen XbaI, XhoI, NotI und DraI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,1 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Die stumpfen Enden des DNA-Fragmentes wurden mit dem Rapid DNA Ligation Kit (Roche

30

## 31

5 Diagnostics, Mannheim) nach Angaben des Herstellers miteinander ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

10 Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK1.

Ausgehend vom Plasmid pWLT1 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden neo1 (SEQ ID NO:55) und neo2 (SEQ ID NO:56) eine Kanamycin-Resistenzcassette amplifiziert.

15 neo1 (SEQ ID NO:55):  
5'-GAGATCTAGACCCGGGGATCCGCTAGCGGGCTGCTAAAGGAAGCGGA-3'

neo2 (SEQ ID NO:56):  
5'-GAGAGGCGCGCCGCTAGCGTGGGCGAAGAACTCCAGCA-3'

20 Neben den zu pWLT1 komplementären Sequenzen, enthält das Oligonukleotid neo1 in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen XbaI, SmaI, BamHI, NheI und das Oligonukleotid neo2 (SEQ ID NO:56) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen AscI und NheI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al.  
25 (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,3 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit den Restriktionsendonukleasen XbaI und AscI (New England Biolabs, Beverly, USA)  
30 geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK1 wurde ebenfalls mit den Restriktionsendonukleasen XbaI und AscI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim) nach Angaben des Herstellers

- dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,1kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben (1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) und Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.
- Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK2.
- Der Vektor pCLiK2 wurde mit der Restriktionsendonuklease DraI (New England Biolabs, Beverly, USA) geschnitten. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde ein ca. 2,3 kb großes Vektorfragment mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers religiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben (1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.
- Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK3.
- Ausgehend vom Plasmid pWLQ2 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden cg1 ((SEQ ID NO:57) und cg2 (SEQ ID NO:58) der Replikationsursprung pHM1519 amplifiziert.

cg1 (SEQ ID NO:57):

5'-GAGAGGGCGGCCGCGCAAAGTCCCGCTTCGTGAA-3'

cg2 (SEQ ID NO:58):

5 5'-GAGAGGGCGGCCGCTCAAGTCGGTCAAGCCACGC-3'

Neben den zu pWLQ2 komplementären Sequenzen, enthalten die Oligonukleotide cg1 (SEQ ID NO:57) und cg2 (SEQ ID NO:58) Schnittstellen für die Restriktionsendonuklease NotI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,7 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease NotI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK3 wurde ebenfalls mit der Restriktionsendonuklease NotI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,3kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5.

Für die Erweiterung von pCLiK5 um eine „multiple cloning site“ (MCS) wurden die beide synthetischen, weitestgehend komplementären Oligonukleotide HS445 ((SEQ ID NO:59) und HS446

(SEQ ID NO:60), die Schnittstellen für die Restriktionsendonukleasen SwaI, XhoI, AatI, ApaI, Asp718, MluI, NdeI, SpeI, EcoRV, Sall, ClaI, BamHI, XbaI und SmaI enthalten, durch gemeinsames Erhitzen auf 95°C und langsames Abkühlen zu einem doppelsträngigen DNA-Fragment vereinigt.

5

HS445 (SEQ ID NO:59):

5'-TCGAATTTAAATCTCGAGAGGCCTGACGTGGGCCCCGGTACCACGCGTCATATGACTAG  
TTCGGACCTAGGGATATCGTCGACATCGATGCTCTTCTGCGTTAATTAACAATTGGGATCC  
TCTAGACCCGGGATTAAAT-3'

10

HS446 (SEQ ID NO:60):

5'-GATCATTTAAATCCCGGGTCTAGAGGATCCCAATTGTTAATTAACGCAGAAGAGCATCGA  
TGTCGACGATATCCCTAGGTCCGAAGTAGTCATATGACGCGTGGTACCGGGCCCCGACGTC  
AGGCCTCTCGAGATTAAAT-3'

15

Der Vektor pCLiK5 wurde mit den Restriktionsendonuklease XhoI und BamHI (New England Biolabs, Beverly, USA) geschnitten und mit alkalischer Phosphatase I (Roche Diagnostics, Mannheim) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 5,0 kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem synthetischen doppelsträngigen DNA-Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

20

25

30

Die Plasmid-DNA eines individuellen Klon wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5MCS.

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Aca-

demy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pCLiK5MCS ist als SEQ ID NO: 63 aufgeführt.

5

**Beispiel 2: Konstruktion von pCLiK5MCS Integrativ sacB**

Ausgehend vom Plasmid pK19mob (Schäfer et al., Gene 145,69-73(1994)) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden BK1732 und BK1733 das *Bacillus subtilis* sacB Gen (kodierend für Levan Sucrase) amplifiziert.

10

BK1732 (SEQ ID NO:61):

5'-GAGAGCGGCCGCGATCCTTTTAAACCCATCAC-3'

15 BK1733 (SEQ ID NO:62):

5'-AGGAGCGGCCGCCATCGGCATTTTCTTTTGCG-3'

Neben den zu pEK19mobsac komplementären Sequenzen, enthalten die Oligonukleotide BK1732 und BK1733 Schnittstellen für die Restriktionsendonuklease NotI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,9 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease NotI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

20

25

Der Vektor pCLiK5MCS (hergestellt gemäß Beispiel 1) wurde ebenfalls mit der Restriktionsendonuklease NotI geschnitten und mit alkalischer Phosphatase I (Roche Diagnostics, Mannheim) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde ein ungefähr 2,4 kb großes Vektorfragment mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers iso-

30



liert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben (1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5MCS integrativ sacB.

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pCLiK5MCS integrativ sacB ist als SEQ ID NO: 64 aufgeführt.

Weitere Vektoren die zur erfindungsgemäßen Expression oder Überproduktion von methH-Genen geeignet sind, können in analoger Weise hergestellt werden.

In den folgenden Beispielen 3 bis 8 wird die schrittweise Konstruktion eines verbesserten Methionin-produzierenden Stammes mit der Bezeichnung LU1479 lysC 311ile ET-16 pC Phsdh meth\_Sc beschrieben.

#### **Beispiel 3: Isolierung des lysC gens aus dem C. glutamicum Stamm LU1479**

Im ersten Schritt der Stammkonstruktion soll ein allelischer Austausch des lysC Wildtypgens, kodierend für das Enzym Aspartatkinase, in C. glutamicum ATCC13032, im folgenden LU1479 genannt, durchgeführt werden. Dabei soll im LysC Gen ein Nukleotidaustausch durchgeführt werden, so dass im resultierenden Protein die Aminosäure Thr an der Position 311 durch die Aminosäure Ile ausgetauscht ist.

Ausgehend von der chromosomalen DNA aus LU1479 als Template für eine PCR Reaktion

wurde mit den Oligonukleotidprimern SEQ ID NO:65 und SEQ ID NO:66 lysC mit Hilfe des Pfu-Turbo PCR Systems (Stratagene USA) nach Angaben des Herstellers amplifiziert. Chromosomale DNA aus *C. glutamicum* ATCC 13032 wurde nach Tauch et al. (1995) Plasmid 33:168-179 oder Eikmanns et al. (1994) Microbiology 140:1817-1828 präpariert. Das amplifizierte Fragment wird an seinem 5'-Ende von einem Sall Restriktionsschnitt und an seinem 3'-Ende von einem MluI Restriktionsschnitt flankiert. Vor der Klonierung wurde das amplifizierte Fragment durch diese beiden Restriktionsenzyme verdaut und mit GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aufgereinigt.

10 SEQ ID NO:65

5'-GAGAGAGAGACGCGTCCCAGTGGCTGAGACGCATC-3'

SEQ ID NO:66

5'-CTCTCTCTGTGCGACGAATTCAATCTTACGGCCTG-3'

15

Das erhaltene Polynukleotid wurde über die Sall und MluI Restriktionsschnitte in pCLIK5 MCS integrativ SacB (im folgenden pCIS genannt; SEQ ID NO: 64 aus Beispiel 2) kloniert und in *E.coli* XL-1 blue transformiert. Eine Selektion auf Plasmid-tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml)-haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht. Das Plasmid wurde isoliert und durch Sequenzierung die erwartete Nukleotidsequenz bestätigt. Die Präparation der Plasmid-DNA wurde nach Methoden und mit Materialien der Firma Quiagen durchgeführt. Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierungsreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet. Das erhaltene Plasmid pCIS lysC ist als SEQ ID NO:77 aufgeführt.

20

25

Die Sequenz SEQ ID NO:77 umfasst die folgenden wesentlichen Teilbereiche:

Position	Art der Sequenz	Beschreibung
155 – 1420	CDS <sup>1)</sup>	lysC
1974 – 2765	CDS	Kanamycin-Resistenz

3032 – 3892 (complement) <sup>2)</sup>	CDS	Replikationsursprung/ E.coli/Plasmid pMB
---	-----	---

<sup>1)</sup> Kodierende Sequenz

<sup>2)</sup> auf Komplementärstrang

#### Beispiel 4: Mutagenese des lysC Gens aus *C. glutamicum*

- 5 Die gerichtete Mutagenese des lysC Gens aus *C. glutamicum* (Beispiel 3) wurde mit dem QuickChange Kit (Fa. Stratagene/USA) nach Angaben des Herstellers durchgeführt. Die Mutagenese wurde im Plasmid pCIS lysC, SEQ ID NO:77 durchgeführt. Für den Austausch von thr311 nach 311ile mit Hilfe der Quickchange Methode (Stratagene) wurden folgende Oligonukleotidprimer synthetisiert

10

SEQ ID NO:67

5'-CGGCACCACCGACATCATCTTCACCTGCCCTCGTTCCG -3'

SEQ ID NO:68

- 15 5'-CGGAACGAGGGCAGGTGAAGATGATGTCGGTGGTGCCG -3'

Der Einsatz dieser Oligonukleotidprimer in der Quickchange Reaktion führt in dem lysC Gen zu einem Austausch des Nukleotids in Position 932 (von C nach T) (vgl. SEQ ID NO:75) und im korrespondierenden Enzym zu einem Aminosäuresubstitution in Position 311 (Thr→Ile) (vgl. SEQ ID NO:76). Der resultierende Aminosäureaustausch Thr311Ile im lysC Gen wurde nach Transformation in *E.coli* XL1-blue und Plasmidpräparation durch Sequenzierung bestätigt. Das Plasmid erhielt die Bezeichnung pCIS lysC thr311ile und ist als SEQ ID NO:78 aufgeführt.

20

Die Sequenz SEQ ID NO:78 umfasst die folgenden wesentlichen Teilbereiche:

25

Position	Art der Sequenz	Beschreibung
155 – 1420	CDS <sup>1)</sup>	lysC mutiert
1974 – 2765	CDS	Kanamycin-Resistenz

3032 – 3892 (complement) <sup>2)</sup>	CDS	Replikationsursprung/ E.coli/Plasmid pMB
---	-----	---

<sup>1)</sup> Kodierende Sequenz

<sup>2)</sup> auf Komplementärstrang

Das Plasmid pCIS lysC thr311ile wurde in *C. glutamicum* LU1479 mittels Elektroporation wie bei  
 5 Liebl, et al. (1989) FEMS Microbiology Letters 53:299-303 beschrieben, transformiert. Modifikationen des Protokolls sind in DE-A-10046870 beschrieben. Die chromosomale Anordnung des lysC-Lokus einzelner Transformanten wurde mit Standardmethoden durch Southernblot und Hybridisierung, wie in Sambrook et al. (1989), Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben, überprüft. Dadurch wurde sichergestellt, dass es sich bei den  
 10 Transformanten um solche handelt, die das transformierte Plasmid durch homologe Rekombination am lysC-Lokus integriert haben. Nach Wachstum solcher Kolonien über Nacht in Medien, die kein Antibiotikum enthielten, wurden die Zellen auf ein Saccharose-CM-Agarmedium (10% Saccharose) ausplattiert und bei 30°C für 24 Stunden inkubiert.

Da das im Vektor pCIS lysC thr311ile enthaltende sacB Gen Saccharose in ein toxisches Produkt umwandelt, können nur solche Kolonien anwachsen, die das sacB Gen durch einen zweiten homologen Rekombinationsschritt zwischen dem Wildtyp lysC Gen und dem mutierten Gen  
 15 lysC thr311ile deletiert haben. Während der homologen Rekombination kann entweder das Wildtyp Gen oder das mutierte Gen zusammen mit dem sacB Gen deletiert werden. Wenn das sacB Gen zusammen mit dem Wildtyp Gen entfernt wird, resultiert eine mutierte Transformante.

20

Anwachsende Kolonien wurden gepickt, und auf einen Kanamycin-sensitiven Phänotyp hin untersucht. Klone mit deletiertem SacB Gen müssen gleichzeitig Kanamycin-sensitives Wachstumsverhalten zeigen. Solche Kan-sensitiven Klone wurde in einem Schüttelkolben auf ihre Lysin-Produktivität hin untersucht (siehe Beispiel 6). Zum Vergleich wurde der nichtbehandelte Stamm  
 25 LU1479 angezogen. Klone mit einer gegenüber der Kontrolle erhöhten Lysin-Produktion wurden selektiert, chromosomale DNA wurde gewonnen und der entsprechende Bereich des lysC Gens wurde durch eine PCR-Reaktion amplifiziert und sequenziert. Ein solcher Klon mit der Eigenschaft erhöhter Lysin-Synthese und nachgewiesener Mutation in lysC an der Stelle 932 wurde mit LU1479 lysC 311ile bezeichnet).

**Beispiel 5: Herstellung Ethionin-resistenter C. glutamicum Stämme**

- Im zweiten Schritt der Stammkonstruktion wurde der erhaltene Stamm LU1479 lysC 311ile (Beispiel 4) behandelt, um eine Ethionin-Resistenz (Kase, H. Nakayama K.Agr. Biol. Chem. 39 153-106 1975 L-methionine production by methionine analog-resistant mutants of Corynebacterium glutamicum) zu induzieren: Eine Übernachtskultur in BHI-Medium (Difco) wurde in Citratpuffer (50mM pH 5,5) gewaschen und bei 30°C für 20 min mit N-Methyl-nitrosoguanidin (10mg/ml in 50mM Citrat pH5,5) behandelt. Nach der Behandlung mit dem chemischen Mutagen N-Methyl-nitrosoguanidin wurden die Zellen gewaschen (Citratpuffer 50mM pH 5,5) und auf ein Medium plattiert, das aus folgenden Komponenten, berechnet auf 500ml, zusammengesetzt war: 10g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.5g KH<sub>2</sub>PO<sub>4</sub>, 0.5g K<sub>2</sub>HPO<sub>4</sub>, 0.125g MgSO<sub>4</sub>·7H<sub>2</sub>O, 21g MOPS, 50mg CaCl<sub>2</sub>, 15mg Proteokatechuat, 0,5mg Biotin, 1mg Thiamin, 5g/l D,L-Ethionin (Sigma Chemicals Deutschland), pH 7,0. Außerdem enthielt das Medium 0.5ml einer Spurensalzlösung aus: 10g/l FeSO<sub>4</sub>·7H<sub>2</sub>O, 1g/l MnSO<sub>4</sub>·H<sub>2</sub>O, 0.1g/l ZnSO<sub>4</sub>·7H<sub>2</sub>O, 0.02g/l CuSO<sub>4</sub>, 0.002g/l NiCl<sub>2</sub>·6H<sub>2</sub>O. Alle Salze wurden in 0,1M HCl gelöst. Das fertig zusammengestellte Medium wurde sterilfiltriert und nach Zugabe von 40ml steriler 50% Glucoselösung, mit flüssigem sterilem Agar in einer Endkonzentration von 1,5% Agar versetzt und in Kulturschalen ausgegossen.
- Auf Platten mit dem beschriebenen Medium wurden mutagenisierte Zellen aufgebracht und 3-7 Tage bei 30°C inkubiert. Erhaltene Klone wurden isoliert, mindestens einmal auf dem Selektionsmedium vereinzelt und dann auf ihre Methionin-Produktivität in einem Schüttelkolben in Medium II untersucht (siehe Beispiel 6)
- Beispiel 6: Herstellung von Methionin mit dem Stamm LU1479 lysC 311ile ET-16.**

Die in Beispiel 5 hergestellten Stämme wurden auf einer Agar-Platte mit CM-Medium für 2 Tag bei 30°C angezogen.

CM-Agar:

- 10,0 g/l D-Glucose, 2,5 g/l NaCl, 2,0 g/l Harnstoff, 10,0 g/l Bacto Pepton (Difco), 5,0 g/l Yeast Extract (Difco), 5,0 g/l Beef Extract (Difco), 22,0 g/l Agar (Difco), autoklaviert (20 min., 121°C)

Anschließend wurden die Zellen von der Platte abgekratzt und in Saline resuspendiert. Für die Hauptkultur wurden 10 ml Medium II und 0,5 g autoklaviertes  $\text{CaCO}_3$  (Riedel de Haen) in einem 100 ml Erlenmeyerkolben mit der Zellsuspension bis zu einer  $\text{OD}_{600\text{nm}}$  von 1,5 beimpft und für  
 5 72h auf einem Orbitalschüttler mit 200 Upm bei 30°C inkubiert.

Medium II:

	40g/l	Saccharose
	60g/l	Melasse (auf 100% Zuckergehalt berechnet)
10	10g/l	$(\text{NH}_4)_2\text{SO}_4$
	0.4g/l	$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$
	0.6g/l	$\text{KH}_2\text{PO}_4$
	0.3mg/l	Thiamin*HCl
15	1mg/l	Biotin (aus einer 1 mg/ml steril filtrierten Stammlösung die mit $\text{NH}_4\text{OH}$ auf pH 8,0 eingestellt wurde)
	2mg/l	$\text{FeSO}_4$
	2mg/l	$\text{MnSO}_4$

mit  $\text{NH}_4\text{OH}$  auf pH 7,8 eingestellt, autoklaviert (121°C, 20 min). Zusätzlich wird Vitamin B12 (Hydroxycobalamin Sigma Chemicals) aus einer Stammlösung (200 µg/ml, steril filtriert) bis zu  
 20 einer Endkonzentration von 100 µg/l zugegeben

Gebildetes Methionin, sowie andere Aminosäuren in der Kulturbrühe wurde mit Hilfe der Aminosäuresäure-Bestimmungsmethode von Agilent auf einer Agilent 1100 Series LC System HPLC. Eine Derivatisierung vor der Säulentrennung mit Ortho-Phthalaldehyd erlaubte die Quan-  
 25 tifizierung der gebildeten Aminosäuren. Die Auftrennung des Aminosäuregemisch fand auf einer Hypersil AA-Säule (Agilent) statt.

Solche Klone wurden isoliert, deren Methionin-Produktivität mindestens doppelt so hoch war, wie die des Ausgangsstamm LU1479 lysC 311ile. Ein solcher Klon wurde für die weiteren Versuche  
 30 eingesetzt und bekam die Bezeichnung LU1479 lysC 311ile ET-16.

**Beispiel 7:** Klonierung von methH aus *Streptomyces coelicolor* und Klonierung in das Plasmid pC

## Phsdh meth\_Sc

- a) Chromosomale DNA wurde aus *Streptomyces coelicolor* Stamm ATCC BAA-471 (von der American Type Strain Culture Collection, (ATCC) Atlanta, USA, unter der Bestellnummer BAA-471D erhältlich) isoliert. Chromosomale DNA aus *C. glutamicum* ATCC 13032 wurde nach Tauch et al. (1995) Plasmid 33:168-179 oder Eikmanns et al. (1994) Microbiology 140:1817-1828 präpariert.

- Mit den Oligonukleotidprimer SEQ ID NO:69 und SEQ ID NO:70, der chromosomalen DNA aus *C. glutamicum* als Template und Pfu Turbo Polymerase (Fa. Stratagene) wurde mit Hilfe der Polymerase-Kettenreaktion (PCR) nach Standardmethoden, wie Innis et al. (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press, ein DNA Fragment von ca. 180 Basenpaaren aus dem nichtkodierenden 5'-Bereich (Promotorregion) der Homoserindehydrogenase (HsDH) amplifiziert. Das amplifizierte Fragment ist an seinem 5'-Ende von einer XhoI-Restriktionsschnittstelle und am 3'-Ende von einem über das Oligo eingeführten zu methH aus *Streptomyces coelicolor* homologen Bereich flankiert.

SEQ ID NO:69

5'-GAGACTCGAGGGAAGGTGAATCGAATTTTCGG-3'

und

SEQ ID NO:70

5'-GTCCCGGGGAGAACGCACGATTCTCCAAAATAATCGC-3'

- Das erhaltene DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

- b) Ausgehend von der chromosomalen DNA aus *Streptomyces coelicolor* als Template für eine PCR Reaktion wurde mit den Oligonukleotidprimern SEQ ID NO:71 und SEQ ID NO:72 ein Teil von methH mit Hilfe des GC-RICH PCR Systems (Roche Diagnostics, Mannheim) nach Angaben des Herstellers amplifiziert. Das amplifizierte Fragment ist an seinem 5'-Ende von einem über das Oligo eingeführten, zur Promotorregion von HsDH aus *C. glutamicum* homologen Bereich flankiert.

SEQ ID NO:71

5'-GAATCGTGCGTTCTCCCCGGGAC -3'

und

SEQ ID NO:72

5 5'-GTAGTTGACCGAGTTGATCACC -3'

Das ca. 1,4 kb große erhaltene DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

- 10 c) In einer weiteren PCR Reaktion wurden die beiden oben erhaltenen Fragmente gemeinsam als Template eingesetzt. Durch die mit dem Oligonukleotidprimern SEQ ID NO:71 und SEQ ID NO:70 eingebrachten, zu dem jeweils anderen Fragment homologen Bereichen, kommt es im Zuge der PCR-Reaktion zu einer Anlagerung beider Fragmente aneinander und einer Verlängerung zu einem durchgehenden DNA-Strang durch die eingesetzte Polymerase. Die Standardmethode wurde dahingehend modifiziert, dass die verwendeten Oligonukleotidprimer SEQ ID NO:69 und SEQ ID NO:72 erst mit Beginn des 2. Zyklus dem Reaktionsansatz zugegeben wurden.

20 Das amplifizierte DNA Fragment von ungefähr 1,6 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit nach Angaben des Herstellers gereinigt. Im Anschluss daran wurde es mit den Restriktionsenzymen XhoI und NotI (Roche Diagnostics, Mannheim) gespalten und gelelektrophoretisch aufgetrennt. Anschließend wurde das ca. 1,6 kb große DNA Fragment mit GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aus der Agarose aufgereinigt.

- 25 d) Der noch fehlende 3'-Bereich von metH wurde ausgehend von der chromosomalen DNA aus *Streptomyces coelicolor* als Template mit den Oligonukleotidprimern SEQ ID NO:73 und SEQ ID NO:74 mit Hilfe des GC-RICH PCR Systems (Roche Diagnostics, Mannheim) nach Angaben des Herstellers amplifiziert. Das amplifizierte Fragment ist an seinem 3'-Ende von einer
- 30 über das Oligo eingeführten EcoRV-Restriktionsschnittstelle flankiert.

SEQ ID NO:73

5'-CCGGCCTGGAGAAGCTCG-3'



und

SEQ ID NO:74

5'-GAGAGATATCCCTCAGCGGGCGTTGAAG-3'

5 Das erhaltene ca. 2,2 kb große DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Im Anschluß daran wurde es mit den Restriktionsenzymen NotI und EcoRV (Roche Diagnostics, Mannheim) gespalten und gelelektrophoretisch aufgetrennt. Anschließend wurde das ca. 2,2 kb große DNA Fragment mit GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aus der Agarose aufgereinigt.

10 e) Der Vektor pClik5MCS SEQ ID NO:63 (Beispiel 1) wurde mit den Restriktionsenzymen XhoI und EcoRV (Roche Diagnostics, Mannheim) geschnitten und ein 5 kb großes Fragment nach elektrophoretischer Auftrennung mit GFX™PCR, DNA and Gel Band Purification Kit isoliert.

15 Das Vektorfragment wurde zusammen mit den beiden geschnittenen und aufgereinigten PCR-Fragmenten mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid-tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

20 Die Präparation der Plasmid DNA wurde nach Methoden und mit Materialien der Fa. Quiagen durchgeführt. Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierungsreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

30 Das entstandene Plasmid pC Phsdh meth<sub>Sc</sub> (*Streptomyces coelicolor*) ist als SEQ ID NO:79 aufgeführt.

Die Sequenz SEQ ID NO:79 umfasst die folgenden wesentlichen Teilbereiche:

Position	Art der Sequenz	Beschreibung
6 - 155	Promotor	HsDH
156 - 3752	CDS <sup>1)</sup>	Meth S. coelicolor
4153 - 4944	CDS	Kanamycin-Resistenz
5211 - 6071 (complement) <sup>2)</sup>	CDS	Replikationsursprung/ E.coli/Plasmid pMB

<sup>1)</sup> Kodierende Sequenz

<sup>2)</sup> auf Komplementärstrang

5

**Beispiel 8:** Transformation des Stammes LU1479 lysC 311ile ET-16 mit dem Plasmid pC Phsdh meth<sub>H</sub>\_Sc

Der Stamm LU1479 lysC 311ile ET-16 (Beispiel 5) wurde mit dem Plasmid pC Phsdh meth<sub>H</sub>\_Sc (Beispiel 7) nach der beschriebenen Methode (Liebl, et al. (1989) FEMS Microbiology Letters 53:299-303) transformiert. Die Transformationsmischung wurde auf CM-Platten plattiert, die zusätzlich 20mg/l Kanamycin enthielten, um eine Selektion auf Plasmid-haltige Zellen zu erreichen. Erhaltene Kan-resistente Klone wurden gepickt und vereinzelt. Die Methionin-Produktivität der Klone wurde in einem Schüttelkolbenversuch (s. Beispiel 6) untersucht. Der Stamm LU1479 lysC 311ile ET-16 pC Phsdh meth<sub>H</sub>\_Sc produzierte im Vergleich zu LU1479 lysC 311ile ET-16 signifikant mehr Methionin.

Patentansprüche

1. Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen  
5 Feinchemikalie, welches folgende Schritte umfasst:
  - a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie  
produzierenden coryneformen Bakterienkultur, wobei in den coryneformen  
Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird,  
welche für ein Protein mit Methionin-Synthase (metF) –Aktivität kodiert;
  - 10 b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den  
Zellen der Bakterien, und
  - c) Isolieren der schwefelhaltigen Feinchemikalie.
2. Verfahren nach Anspruch 1, wobei die schwefelhaltige Feinchemikalie L-Methionin  
15 umfasst.
3. Verfahren nach einem der vorhergehenden Ansprüche, wobei sich die heterologe metF-  
kodierende Nukleotidsequenz zur metF-kodierenden Sequenz aus *Corynebacterium*  
*glutamicum* ATCC 13032 eine Sequenzhomologie vom weniger als 100% aufweist.
- 20 4. Verfahren nach Anspruch 3, wobei die metF-kodierende Sequenz aus einem der  
folgenden Organismen abgeleitet ist:

Organismus	Stammsammlung
<i>Corynebacterium diptheriae</i>	ATCC 14779
<i>Streptomyces lividans</i>	ATCC 19844
<i>Streptomyces coelicolor</i>	ATCC 10147
<i>Aquifex aeolicus</i>	DSM 6858
<i>Burkholderia cepacia</i>	ATCC 25416
<i>Nitrosomonas europaea</i>	ATCC 19718
<i>Pseudomonas aeruginosa</i>	ATCC 17933
<i>Xylella fastidiosa</i>	ATCC 35881
<i>Pseudomonas fluorescens</i>	ATCC 13525
<i>Schizosaccharomyces pombe</i>	ATCC 24969
<i>Saccharomyces cerevisiae</i>	ATCC 10751
<i>Erwinia carotovora</i>	ATCC 15713
<i>Klebsiella pneumoniae</i>	ATCC 700721
<i>Salmonella typhi</i>	ATCC 12839
<i>Salmonella typhimurium</i>	ATCC 15277
<i>Escherichia coli</i> K12	ATCC55151

<i>Vibrio cholerae</i>	ATCC 39315
<i>Haemophilus influenzae</i>	ATCC 51907
<i>Caulobacter crescentus</i>	ATCC 19089
<i>Actinobacillus actinomycetemcomitans</i>	ATCC 33384
<i>Neisseria meningitis</i>	ATCC 6253
<i>Rhodobacter capsulatus</i>	ATCC 11166
<i>Campylobacter jejuni</i>	ATCC 33560
<i>Lactococcus lactis</i>	ATCC 7962
<i>Prochlorococcus marinus</i>	PCC7118
<i>Bacillus stearothermophilus</i>	ATCC 12980

5. Verfahren nach einem der vorhergehenden Ansprüche, wobei die metF-kodierende Sequenz eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit metF-Aktivität kodiert, umfasst.
6. Verfahren nach einem der vorhergehenden Ansprüche, wobei die metF-kodierende Sequenz für ein Protein mit metF-Aktivität kodiert, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 und 54 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metF-Aktivität steht, umfasst.
7. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metF-Sequenz eine in coryneformen Bakterien replizierbare oder eine stabil in das Chromosom integrierte DNA oder eine RNA ist.
8. Verfahren gemäß Anspruch 7, wobei man
  - a) einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden metF-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
  - b) einen Stamm einsetzt, in dem die kodierende metF-Sequenz in das Chromosom des Bakteriums integriert wurde
9. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metF-Sequenz überexprimiert wird.

10. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie verstärkt ist oder derart mutiert ist, dass es durch Stoffwechselmetabolite nicht in seiner Aktivität beeinflusst wird.
11. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet sind, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.
12. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter
- a) dem für eine Aspartatkinase kodierenden Gen *lysC*,
  - b) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen *gap*,
  - c) dem für die 3-Phosphoglycerat Kinase kodierenden Gen *pgk*,
  - d) dem für die Pyruvat Carboxylase kodierenden Gen *pyc*,
  - e) dem für die Triosephosphat Isomerase kodierenden Gen *tpl*,
  - f) dem für die Homoserin O-Acetyltransferase kodierenden Gen *metA*,
  - g) dem für die Cystathionin-gamma-Synthase kodierenden Gen *metB*,
  - h) dem für die Cystathionin-gamma-Lyase kodierenden Gen *metC*,
  - i) dem für die Serin-Hydroxymethyltransferase kodierenden Gen *glyA*,
  - j) dem für die O-Acetylhomoserin-Sulphydrylase kodierenden Gen *metY*,
  - k) dem für das *methH* Gen, das für die Vitamin B12 abhängige Methionin-Synthase kodiert,
  - l) dem für das *serC* Gen, das für die Phosphoserin-Aminotransferase kodiert,
  - m) dem *serB* Gen, das für die Phosphoserin-Phosphatase kodiert,
  - n) dem *cysE* Gen, das für die Serine Acetyl-Transferase kodiert, und
  - o) dem *hom* Gen, das eine Homoserin-Dehydrogenase kodiert,
- überexprimiert oder so mutiert ist, dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden.

13. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneformen Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter
- a) dem für die Homoserine-Kinase kodierenden Gen *thrB*,
  - b) dem für die Threonin Dehydratase kodierenden Gen *ilvA*,
  - 5 c) dem für die Threonin Synthase kodierenden Gen *thrC*
  - d) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen *ddh*
  - e) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen *pck*,
  - f) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen *pgi*,
  - g) dem für die Pyruvat-Oxidase kodierenden Gen *poxB*,
  - 10 h) dem für die Dihydrodipicolinat Synthase kodierenden Gen *dapA*,
  - i) dem für die Dihydrodipicolinat Reduktase kodierenden Gen *dapB*; oder
  - j) dem für die Diaminopicolinat Decarboxylase kodierenden Gen
- durch Veränderung der Expressionsrate oder durch Einführung einer gezielten Mutation
- 15 abschwächt ist.
14. Verfahren gemäß einem oder mehreren der vorhergehenden Ansprüche, wobei man Mikroorganismen der Art *Corynebacterium glutamicum* einsetzt.
- 20 15. Verfahren zur Herstellung eines L-Methionin haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst
- a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
  - b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
  - 25 c) Entfernung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
  - d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.
- 30 16. Verfahren gemäß Anspruch 15, wobei man Mikroorganismen gemäß der Definition in einem der Ansprüche 1 bis 14 einsetzt.

## SEQUENZPROTOKOLL

&lt;110&gt; BASF Aktiengesellschaft

&lt;120&gt; Meth

&lt;130&gt; M/43120

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 79

&lt;210&gt; 1

&lt;211&gt; 3597

&lt;212&gt; DNA

&lt;213&gt; Streptomyces coelicolor

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (3594)

&lt;223&gt; RSX14254

&lt;400&gt; 1

gtg	cgt	tct	ccc	cgg	gac	gtc	cca	cga	cgg	gcg	gca	ccg	ggc	aga	ggc	48
Val	Arg	Ser	Pro	Arg	Asp	Val	Pro	Arg	Arg	Ala	Ala	Pro	Gly	Arg	Gly	
1				5					10					15		

aaa	gcc	gac	agc	cgt	cgc	atc	cta	ggg	agc	cct	ttc	atg	gcc	tcg	tcg	96
Lys	Ala	Asp	Ser	Arg	Arg	Ile	Leu	Gly	Ser	Pro	Phe	Met	Ala	Ser	Ser	
			20					25					30			

cca	tcc	acc	ccg	ccc	gcc	gac	acc	cgc	acc	cgc	gtg	tcc	gcc	ctc	cga	144
Pro	Ser	Thr	Pro	Pro	Ala	Asp	Thr	Arg	Thr	Arg	Val	Ser	Ala	Leu	Arg	
		35					40					45				

gag	gcc	ctc	gcc	acc	cgc	gtg	gtg	gtc	gcc	gac	ggc	gcc	atg	ggc	acc	192
Glu	Ala	Leu	Ala	Thr	Arg	Val	Val	Val	Ala	Asp	Gly	Ala	Met	Gly	Thr	
	50					55					60					

atg	ctc	cag	gcc	cag	aac	ccc	acg	ctg	gac	gac	ttc	cag	cag	ctc	gaa	240
Met	Leu	Gln	Ala	Gln	Asn	Pro	Thr	Leu	Asp	Asp	Phe	Gln	Gln	Leu	Glu	
	65				70					75					80	

ggg	tgc	aac	gag	gtc	ctg	aac	ctc	acc	cgg	ccc	gac	atc	gtc	cgc	tcg	288
Gly	Cys	Asn	Glu	Val	Leu	Asn	Leu	Thr	Arg	Pro	Asp	Ile	Val	Arg	Ser	
			85					90						95		

gtg	cac	gag	gag	tac	ttc	gcg	gcc	ggc	gtc	gac	tgc	gtc	gag	acc	aac	336
Val	His	Glu	Glu	Tyr	Phe	Ala	Ala	Gly	Val	Asp	Cys	Val	Glu	Thr	Asn	
		100						105					110			

acc	ttc	ggc	gcc	aac	cac	tcc	gcc	ctg	ggc	gag	tac	gac	atc	ccc	gag	384
Thr	Phe	Gly	Ala	Asn	His	Ser	Ala	Leu	Gly	Glu	Tyr	Asp	Ile	Pro	Glu	
		115					120					125				

cgc	gtc	cac	gaa	ctg	tcc	gag	gcc	ggc	gcc	cgc	gtc	gcc	cgc	gag	gtc	432
Arg	Val	His	Glu	Leu	Ser	Glu	Ala	Gly	Ala	Arg	Val	Ala	Arg	Glu	Val	
	130					135					140					

gcc	gac	gag	ttc	ggc	gcc	cgc	gac	ggc	cgg	cag	cgc	tgg	gtg	ctg	ggc	480
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

2

Ala 145	Asp	Glu	Phe	Gly	Ala 150	Arg	Asp	Gly	Arg	Gln 155	Arg	Trp	Val	Leu 160	Gly	
tcc Ser	atg Met	ggc Gly	ccc Pro	ggc Gly	acc Thr	aag Lys	ctc Leu	ccc Pro	acc Thr	ctc Leu	ggc Gly	cac His	gcc Ala	ccg Pro	tac Tyr	528
acc Thr	gtc Val	ctg Leu	cgc Arg	gac Asp	gcc Ala	tac Tyr	cag Gln	cgc Arg	aac Asn	gcc Ala	gag Glu	gga Gly	ctg Leu	gtc Val	gcg Ala	576
ggc Gly	ggc Gly	gcg Ala	gac Asp	gca Ala	ctg Leu	ctg Leu	gtg Val	gag Glu	acc Thr	acg Thr	cag Gln	gac Asp	ctg Leu	ctc Leu	cag Gln	624
acc Thr	aag Lys	gcc Ala	tcg Ser	gtg Val	ctc Leu	ggc Gly	gcc Ala	cgg Arg	cgc Arg	gcc Ala	ctg Leu	gac Asp	gtc Val	ctc Leu	ggc Gly	672
ctc Leu	gac Asp	ctg Leu	ccg Pro	ctc Leu	atc Ile	gtg Val	tcc Ser	gtc Val	acc Thr	gtc Val	gag Glu	acc Thr	acc Thr	ggc Gly	acc Thr	720
atg Met	ctg Leu	ctc Leu	ggc Gly	tcg Ser	gag Glu	atc Ile	ggc Gly	gcc Ala	gcg Ala	ctc Leu	acc Thr	gcg Ala	ctg Leu	gaa Glu	ccg Pro	768
ctc Leu	ggc Gly	atc Ile	gac Asp	atg Met	atc Ile	ggc Gly	ctg Leu	aac Asn	tgc Cys	gcc Ala	acc Thr	ggc Gly	ccc Pro	gcc Ala	gag Glu	816
atg Met	agc Ser	gag Glu	cac His	ctg Leu	cgc Arg	tac Tyr	ctc Leu	gcc Ala	cgg Arg	cac His	tcc Ser	cgc Arg	atc Ile	ccg Pro	ctg Leu	864
acc Thr	tgc Cys	atg Met	ccc Pro	aac Asn	gcc Ala	ggt Gly	ctg Leu	ccc Pro	gtc Val	ctc Leu	ggc Gly	aag Lys	gac Asp	ggc Gly	gcc Ala	912
cac His	tac Tyr	ccg Pro	ctg Leu	acc Thr	gcg Ala	ccc Pro	gag Glu	ctg Leu	gcc Ala	gac Asp	gca Ala	cac His	gag Glu	acc Thr	ttc Phe	960
gtg Val	cgc Arg	gag Glu	tac Tyr	ggc Gly	ctg Leu	tcc Ser	ctg Leu	gtc Val	ggc Gly	ggc Gly	tgc Cys	tgc Cys	ggc Gly	acc Thr	acg Thr	1008
ccc Pro	gag Glu	cac His	ctg Leu	cgc Arg	cag Gln	gtc Val	gtc Val	gag Glu	cgg Arg	gtc Val	cgg Arg	gac Asp	acc Thr	gcc Ala	ccc Pro	1056
acc Thr	gca Ala	cgc Arg	gac Asp	ccg Pro	cgc Arg	ccc Pro	gag Glu	ccc Pro	ggc Gly	gcc Ala	gcc Ala	tcg Ser	ctc Leu	tac Tyr	cag Gln	1104
acc Thr	gtg Val	ccc Pro	ttc Phe	cgc Arg	cag Gln	gac Asp	acc Thr	tcc Ser	tac Tyr	ctg Leu	gcc Ala	atc Ile	ggc Gly	gag Glu	cgc Arg	1152
acc Thr	aac Asn	gcc Ala	aac Asn	ggg Gly	tcc Ser	aag Lys	aag Lys	ttc Phe	cgc Arg	gag Glu	gcc Ala	atg Met	ctg Leu	gac Asp	ggc Gly	1200



cgc tgg gac gac tgc gtc gag atg gcc cgc gac cag atc cgc gaa ggc Arg Trp Asp Asp Cys Val Glu Met Ala Arg Asp Gln Ile Arg Glu Gly 405 410 415	1248
gcg cac atg ctc gac ctc tgc gtc gac tac gtc ggc cgg gac ggc gtc Ala His Met Leu Asp Leu Cys Val Asp Tyr Val Gly Arg Asp Gly Val 420 425 430	1296
gcc gac atg gag gaa ctg gcc ggc cgg ttc gcc acc gcc tcc acg ctg Ala Asp Met Glu Glu Leu Ala Gly Arg Phe Ala Thr Ala Ser Thr Leu 435 440 445	1344
ccg atc gtc ctc gac tcc acc gag gtc gac gtc atc cgg gcc ggc ctg Pro Ile Val Leu Asp Ser Thr Glu Val Asp Val Ile Arg Ala Gly Leu 450 455 460	1392
gag aag ctc ggc ggc cgc gcg gtg atc aac tcg gtc aac tac gag gac Glu Lys Leu Gly Gly Arg Ala Val Ile Asn Ser Val Asn Tyr Glu Asp 465 470 475 480	1440
ggc gcc ggc ccc gag tcc cgg ttc gcc cgc gtc acg aag ctc gcc cgg Gly Ala Gly Pro Glu Ser Arg Phe Ala Arg Val Thr Lys Leu Ala Arg 485 490 495	1488
gag cac ggc gcc gcg ctg atc gcg ctg acc atc gac gag gtg gga cag Glu His Gly Ala Ala Leu Ile Ala Leu Thr Ile Asp Glu Val Gly Gln 500 505 510	1536
gcc cgc acc gcc gag aag aag gtc gag atc gcc gaa cgg ctc atc gac Ala Arg Thr Ala Glu Lys Lys Val Glu Ile Ala Glu Arg Leu Ile Asp 515 520 525	1584
gac ctc acc ggc aac tgg ggc atc cac gag tcc gac atc ctc gtc gac Asp Leu Thr Gly Asn Trp Gly Ile His Glu Ser Asp Ile Leu Val Asp 530 535 540	1632
tgc ctg acc ttc acc atc tgc acc ggc cag gag gag tcc cgc aag gac Cys Leu Thr Phe Thr Ile Cys Thr Gly Gln Glu Glu Ser Arg Lys Asp 545 550 555 560	1680
ggc ctg gcc acc atc gag ggc atc cgg gaa ctc aag cgg cgc cac ccg Gly Leu Ala Thr Ile Glu Gly Ile Arg Glu Leu Lys Arg Arg His Pro 565 570 575	1728
gac gtg cag acc acg ctc ggc ctg tcg aac atc tcc ttc ggc ctc aac Asp Val Gln Thr Thr Leu Gly Leu Ser Asn Ile Ser Phe Gly Leu Asn 580 585 590	1776
ccg gcc gcc cgc atc ctg ctc aac tcc gtc ttc ctc gac gaa tgc gtc Pro Ala Ala Arg Ile Leu Leu Asn Ser Val Phe Leu Asp Glu Cys Val 595 600 605	1824
aag gcc ggc ctg gac tcg gcc atc gtg cac gcg agc aag atc ctg ccg Lys Ala Gly Leu Asp Ser Ala Ile Val His Ala Ser Lys Ile Leu Pro 610 615 620	1872
atc gcc cgc ttc gac gag gag cag gtc acc acc gcc ctc gac ttg atc Ile Ala Arg Phe Asp Glu Glu Gln Val Thr Thr Ala Leu Asp Leu Ile 625 630 635 640	1920
tac gac cgc cgc cgc gag ggc tac gac ccc ctg caa aag ctc atg cag	1968

4

Tyr Asp Arg Arg Arg Glu Gly Tyr Asp Pro Leu Gln Lys Leu Met Gln	
645 650 655	
ctc ttc gag ggc gcc acc gcc aag tcg ctg aag gcc tcc aag gcc gag	2016
Leu Phe Glu Gly Ala Thr Ala Lys Ser Leu Lys Ala Ser Lys Ala Glu	
660 665 670	
gaa ctg gcc gcc ctc ccg ctg gag gag cgc ctc aag cgc cgc atc atc	2064
Glu Leu Ala Ala Leu Pro Leu Glu Glu Arg Leu Lys Arg Arg Ile Ile	
675 680 685	
gac ggc gag aag aac ggc ctc gaa cag gac ctc gac gag gcc ctc cgg	2112
Asp Gly Glu Lys Asn Gly Leu Glu Gln Asp Leu Asp Glu Ala Leu Arg	
690 695 700	
gag cgc ccg gcc ctc gag atc gtc aac gac acc ctg ctc gac ggt atg	2160
Glu Arg Pro Ala Leu Glu Ile Val Asn Asp Thr Leu Leu Asp Gly Met	
705 710 715 720	
aag gtc gtc ggc gag ctg ttc ggc tcc ggc cag atg cag ctg ccg ttc	2208
Lys Val Val Gly Glu Leu Phe Gly Ser Gly Gln Met Gln Leu Pro Phe	
725 730 735	
gtg ctc cag tcc gcc gag gtc atg aag acc gcg gtg gcc cac ctg gag	2256
Val Leu Gln Ser Ala Glu Val Met Lys Thr Ala Val Ala His Leu Glu	
740 745 750	
ccg cac atg gag aag acc gac gac gac ggc aag ggc acg atc gtg ctg	2304
Pro His Met Glu Lys Thr Asp Asp Asp Gly Lys Gly Thr Ile Val Leu	
755 760 765	
gcc acc gtc cgc ggc gac gtc cac gac atc ggc aag aac ctc gtc gac	2352
Ala Thr Val Arg Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Asp	
770 775 780	
atc atc ctg tcc aac aac ggc tac aac gtc gtc aac ctc ggc atc aag	2400
Ile Ile Leu Ser Asn Asn Gly Tyr Asn Val Val Asn Leu Gly Ile Lys	
785 790 795 800	
cag ccc gtc tcc gcg atc ctg gaa gcg gcc gac gag cac cgg gcc gac	2448
Gln Pro Val Ser Ala Ile Leu Glu Ala Ala Asp Glu His Arg Ala Asp	
805 810 815	
gtc atc ggc atg tcc ggc ctc ctc gtc aag tcc acg gtg atc atg aag	2496
Val Ile Gly Met Ser Gly Leu Leu Val Lys Ser Thr Val Ile Met Lys	
820 825 830	
gag aac ctg gag gag ctg aac cag cgc aag ctg gcc gcc gac tac ccg	2544
Glu Asn Leu Glu Glu Leu Asn Gln Arg Lys Leu Ala Ala Asp Tyr Pro	
835 840 845	
gtc atc ctc ggc ggc gcc gcc ctc acc agg gcc tac gtc gaa cag gac	2592
Val Ile Leu Gly Gly Ala Ala Leu Thr Arg Ala Tyr Val Glu Gln Asp	
850 855 860	
ctg cac gag atc tac gac ggc gag gtc cgc tac gcc cgc gac gcc ttc	2640
Leu His Glu Ile Tyr Asp Gly Glu Val Arg Tyr Ala Arg Asp Ala Phe	
865 870 875 880	
gag ggc ctg cgc ctc atg gac gcc ctc atc ggc atc aag cgc ggc gtg	2688
Glu Gly Leu Arg Leu Met Asp Ala Leu Ile Gly Ile Lys Arg Gly Val	
885 890 895	

ccc ggc gcc aag ctg ccg gag ctg aag cag cgc cgg gtg cgg gcc gcc	2736
Pro Gly Ala Lys Leu Pro Glu Leu Lys Gln Arg Arg Val Arg Ala Ala	
900 905 910	
acc gtc gag atc gac gag cgc ccc gag gaa ggc cac gtc cgc tcc gac	2784
Thr Val Glu Ile Asp Glu Arg Pro Glu Glu Gly His Val Arg Ser Asp	
915 920 925	
gtc gcc acc gac aac ccg gtc ccg acc ccg ccc ttc cgc ggc acc cgc	2832
Val Ala Thr Asp Asn Pro Val Pro Thr Pro Phe Arg Gly Thr Arg	
930 935 940	
gtc gtc aag ggc atc cag ctc aag gag tac gcc tcc tgg ctc gac gag	2880
Val Val Lys Gly Ile Gln Leu Lys Glu Tyr Ala Ser Trp Leu Asp Glu	
945 950 955 960	
ggc gcc ctc ttc aag ggc cag tgg ggc ctc aag cag gcc cgc acc ggc	2928
Gly Ala Leu Phe Lys Gly Gln Trp Gly Leu Lys Gln Ala Arg Thr Gly	
965 970 975	
gag gga ccc tcc tac gag gaa ctg gtc gag tcc gag ggc cgg ccg cgg	2976
Glu Gly Pro Ser Tyr Glu Glu Leu Val Glu Ser Glu Gly Arg Pro Arg	
980 985 990	
ctg cgc ggc ctg ctc gac cgg ctc cag acg gac aac ctt ttg gag gcg	3024
Leu Arg Gly Leu Leu Asp Arg Leu Gln Thr Asp Asn Leu Leu Glu Ala	
995 1000 1005	
gcc gtg gtc tac ggc tac ttc ccc tgc gtc tcc aag gac gac gac ctg	3072
Ala Val Val Tyr Gly Tyr Phe Pro Cys Val Ser Lys Asp Asp Asp Leu	
1010 1015 1020	
atc gtc ctc gac gac gac ggc aac gaa cgc acc cgc ttc acc ttc ccc	3120
Ile Val Leu Asp Asp Asp Gly Asn Glu Arg Thr Arg Phe Thr Phe Pro	
1025 1030 1035 1040	
cgc cag cgc cgc ggc cgg cgc ctg tgc ctg gcc gac ttc ttc cgc ccg	3168
Arg Gln Arg Arg Gly Arg Arg Leu Cys Leu Ala Asp Phe Phe Arg Pro	
1045 1050 1055	
gag gag tcc ggc gag acc gac gtg gtc ggc ttc cag gtc gtc acc gtc	3216
Glu Glu Ser Gly Glu Thr Asp Val Val Gly Phe Gln Val Val Thr Val	
1060 1065 1070	
ggc tcc cgc atc ggc gag gag acg gcc cgc atg ttc gag gcc aac gcc	3264
Gly Ser Arg Ile Gly Glu Glu Thr Ala Arg Met Phe Glu Ala Asn Ala	
1075 1080 1085	
tac cgc gac tat ctc gag ctg cac ggc ctg tcc gtg cag ctc gcc gag	3312
Tyr Arg Asp Tyr Leu Glu Leu His Gly Leu Ser Val Gln Leu Ala Glu	
1090 1095 1100	
gcc ctc gcc gag tac tgg cac gcg cgc gtg cgc tcg gaa ctc ggc ttc	3360
Ala Leu Ala Glu Tyr Trp His Ala Arg Val Arg Ser Glu Leu Gly Phe	
1105 1110 1115 1120	
gcc ggg gag gac ccg gcc gag atg gag gac atg ttc gcc ctg aag tac	3408
Ala Gly Glu Asp Pro Ala Glu Met Glu Asp Met Phe Ala Leu Lys Tyr	
1125 1130 1135	

6

cgg ggt gcc cgc ttc tcc ctc ggc tac ggc gcc tgc ccc gac ctg gag 3456  
 Arg Gly Ala Arg Phe Ser Leu Gly Tyr Gly Ala Cys Pro Asp Leu Glu  
 1140 1145 1150

gac cgc gcc aag atc gcc gcc ctg ctg gag ccc gag cgc atc ggc gtc 3504  
 Asp Arg Ala Lys Ile Ala Ala Leu Leu Glu Pro Glu Arg Ile Gly Val  
 1155 1160 1165

cac cta tcc gag gag ttc cag ctc cac ccc gag cag tcc acc gac gcc 3552  
 His Leu Ser Glu Glu Phe Gln Leu His Pro Glu Gln Ser Thr Asp Ala  
 1170 1175 1180

atc gtc atc cac cac ccg gag gcc aag tac ttc aac gcc cgc 3594  
 Ile Val Ile His His Pro Glu Ala Lys Tyr Phe Asn Ala Arg  
 1185 1190 1195

tga 3597

<210> 2

<211> 1198

<212> PRT

<213> Streptomyces coelicolor

<400> 2

Val Arg Ser Pro Arg Asp Val Pro Arg Arg Ala Ala Pro Gly Arg Gly  
 1 5 10 15

Lys Ala Asp Ser Arg Arg Ile Leu Gly Ser Pro Phe Met Ala Ser Ser  
 20 25 30

Pro Ser Thr Pro Pro Ala Asp Thr Arg Thr Arg Val Ser Ala Leu Arg  
 35 40 45

Glu Ala Leu Ala Thr Arg Val Val Val Ala Asp Gly Ala Met Gly Thr  
 50 55 60

Met Leu Gln Ala Gln Asn Pro Thr Leu Asp Asp Phe Gln Gln Leu Glu  
 65 70 75 80

Gly Cys Asn Glu Val Leu Asn Leu Thr Arg Pro Asp Ile Val Arg Ser  
 85 90 95

Val His Glu Glu Tyr Phe Ala Ala Gly Val Asp Cys Val Glu Thr Asn  
 100 105 110

Thr Phe Gly Ala Asn His Ser Ala Leu Gly Glu Tyr Asp Ile Pro Glu  
 115 120 125

Arg Val His Glu Leu Ser Glu Ala Gly Ala Arg Val Ala Arg Glu Val  
 130 135 140

Ala Asp Glu Phe Gly Ala Arg Asp Gly Arg Gln Arg Trp Val Leu Gly  
 145 150 155 160

Ser Met Gly Pro Gly Thr Lys Leu Pro Thr Leu Gly His Ala Pro Tyr  
 165 170 175

Thr Val Leu Arg Asp Ala Tyr Gln Arg Asn Ala Glu Gly Leu Val Ala  
 180 185 190

Gly Gly Ala Asp Ala Leu Leu Val Glu Thr Thr Gln Asp Leu Leu Gln

195	200	205
Thr Lys Ala Ser Val Leu Gly Ala Arg Arg Ala Leu Asp Val Leu Gly 210 215 220		
Leu Asp Leu Pro Leu Ile Val Ser Val Thr Val Glu Thr Thr Gly Thr 225 230 235 240		
Met Leu Leu Gly Ser Glu Ile Gly Ala Ala Leu Thr Ala Leu Glu Pro 245 250 255		
Leu Gly Ile Asp Met Ile Gly Leu Asn Cys Ala Thr Gly Pro Ala Glu 260 265 270		
Met Ser Glu His Leu Arg Tyr Leu Ala Arg His Ser Arg Ile Pro Leu 275 280 285		
Thr Cys Met Pro Asn Ala Gly Leu Pro Val Leu Gly Lys Asp Gly Ala 290 295 300		
His Tyr Pro Leu Thr Ala Pro Glu Leu Ala Asp Ala His Glu Thr Phe 305 310 315 320		
Val Arg Glu Tyr Gly Leu Ser Leu Val Gly Gly Cys Cys Gly Thr Thr 325 330 335		
Pro Glu His Leu Arg Gln Val Val Glu Arg Val Arg Asp Thr Ala Pro 340 345 350		
Thr Ala Arg Asp Pro Arg Pro Glu Pro Gly Ala Ala Ser Leu Tyr Gln 355 360 365		
Thr Val Pro Phe Arg Gln Asp Thr Ser Tyr Leu Ala Ile Gly Glu Arg 370 375 380		
Thr Asn Ala Asn Gly Ser Lys Lys Phe Arg Glu Ala Met Leu Asp Gly 385 390 395 400		
Arg Trp Asp Asp Cys Val Glu Met Ala Arg Asp Gln Ile Arg Glu Gly 405 410 415		
Ala His Met Leu Asp Leu Cys Val Asp Tyr Val Gly Arg Asp Gly Val 420 425 430		
Ala Asp Met Glu Glu Leu Ala Gly Arg Phe Ala Thr Ala Ser Thr Leu 435 440 445		
Pro Ile Val Leu Asp Ser Thr Glu Val Asp Val Ile Arg Ala Gly Leu 450 455 460		
Glu Lys Leu Gly Gly Arg Ala Val Ile Asn Ser Val Asn Tyr Glu Asp 465 470 475 480		
Gly Ala Gly Pro Glu Ser Arg Phe Ala Arg Val Thr Lys Leu Ala Arg 485 490 495		
Glu His Gly Ala Ala Leu Ile Ala Leu Thr Ile Asp Glu Val Gly Gln 500 505 510		
Ala Arg Thr Ala Glu Lys Lys Val Glu Ile Ala Glu Arg Leu Ile Asp 515 520 525		

8

Asp Leu Thr Gly Asn Trp Gly Ile His Glu Ser Asp Ile Leu Val Asp  
 530 535 540

Cys Leu Thr Phe Thr Ile Cys Thr Gly Gln Glu Glu Ser Arg Lys Asp  
 545 550 555 560

Gly Leu Ala Thr Ile Glu Gly Ile Arg Glu Leu Lys Arg Arg His Pro  
 565 570 575

Asp Val Gln Thr Thr Leu Gly Leu Ser Asn Ile Ser Phe Gly Leu Asn  
 580 585 590

Pro Ala Ala Arg Ile Leu Leu Asn Ser Val Phe Leu Asp Glu Cys Val  
 595 600 605

Lys Ala Gly Leu Asp Ser Ala Ile Val His Ala Ser Lys Ile Leu Pro  
 610 615 620

Ile Ala Arg Phe Asp Glu Glu Gln Val Thr Thr Ala Leu Asp Leu Ile  
 625 630 635 640

Tyr Asp Arg Arg Arg Glu Gly Tyr Asp Pro Leu Gln Lys Leu Met Gln  
 645 650 655

Leu Phe Glu Gly Ala Thr Ala Lys Ser Leu Lys Ala Ser Lys Ala Glu  
 660 665 670

Glu Leu Ala Ala Leu Pro Leu Glu Glu Arg Leu Lys Arg Arg Ile Ile  
 675 680 685

Asp Gly Glu Lys Asn Gly Leu Glu Gln Asp Leu Asp Glu Ala Leu Arg  
 690 695 700

Glu Arg Pro Ala Leu Glu Ile Val Asn Asp Thr Leu Leu Asp Gly Met  
 705 710 715 720

Lys Val Val Gly Glu Leu Phe Gly Ser Gly Gln Met Gln Leu Pro Phe  
 725 730 735

Val Leu Gln Ser Ala Glu Val Met Lys Thr Ala Val Ala His Leu Glu  
 740 745 750

Pro His Met Glu Lys Thr Asp Asp Asp Gly Lys Gly Thr Ile Val Leu  
 755 760 765

Ala Thr Val Arg Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Asp  
 770 775 780

Ile Ile Leu Ser Asn Asn Gly Tyr Asn Val Val Asn Leu Gly Ile Lys  
 785 790 795 800

Gln Pro Val Ser Ala Ile Leu Glu Ala Ala Asp Glu His Arg Ala Asp  
 805 810 815

Val Ile Gly Met Ser Gly Leu Leu Val Lys Ser Thr Val Ile Met Lys  
 820 825 830

Glu Asn Leu Glu Glu Leu Asn Gln Arg Lys Leu Ala Ala Asp Tyr Pro  
 835 840 845

Val Ile Leu Gly Gly Ala Ala Leu Thr Arg Ala Tyr Val Glu Gln Asp  
 850 855 860

Leu His Glu Ile Tyr Asp Gly Glu Val Arg Tyr Ala Arg Asp Ala Phe  
 865 870 875 880  
 Glu Gly Leu Arg Leu Met Asp Ala Leu Ile Gly Ile Lys Arg Gly Val  
 885 890 895  
 Pro Gly Ala Lys Leu Pro Glu Leu Lys Gln Arg Arg Val Arg Ala Ala  
 900 905 910  
 Thr Val Glu Ile Asp Glu Arg Pro Glu Glu Gly His Val Arg Ser Asp  
 915 920 925  
 Val Ala Thr Asp Asn Pro Val Pro Thr Pro Pro Phe Arg Gly Thr Arg  
 930 935 940  
 Val Val Lys Gly Ile Gln Leu Lys Glu Tyr Ala Ser Trp Leu Asp Glu  
 945 950 955 960  
 Gly Ala Leu Phe Lys Gly Gln Trp Gly Leu Lys Gln Ala Arg Thr Gly  
 965 970 975  
 Glu Gly Pro Ser Tyr Glu Glu Leu Val Glu Ser Glu Gly Arg Pro Arg  
 980 985 990  
 Leu Arg Gly Leu Leu Asp Arg Leu Gln Thr Asp Asn Leu Leu Glu Ala  
 995 1000 1005  
 Ala Val Val Tyr Gly Tyr Phe Pro Cys Val Ser Lys Asp Asp Asp Leu  
 1010 1015 1020  
 Ile Val Leu Asp Asp Asp Gly Asn Glu Arg Thr Arg Phe Thr Phe Pro  
 1025 1030 1035 1040  
 Arg Gln Arg Arg Gly Arg Arg Leu Cys Leu Ala Asp Phe Phe Arg Pro  
 1045 1050 1055  
 Glu Glu Ser Gly Glu Thr Asp Val Val Gly Phe Gln Val Val Thr Val  
 1060 1065 1070  
 Gly Ser Arg Ile Gly Glu Glu Thr Ala Arg Met Phe Glu Ala Asn Ala  
 1075 1080 1085  
 Tyr Arg Asp Tyr Leu Glu Leu His Gly Leu Ser Val Gln Leu Ala Glu  
 1090 1095 1100  
 Ala Leu Ala Glu Tyr Trp His Ala Arg Val Arg Ser Glu Leu Gly Phe  
 1105 1110 1115 1120  
 Ala Gly Glu Asp Pro Ala Glu Met Glu Asp Met Phe Ala Leu Lys Tyr  
 1125 1130 1135  
 Arg Gly Ala Arg Phe Ser Leu Gly Tyr Gly Ala Cys Pro Asp Leu Glu  
 1140 1145 1150  
 Asp Arg Ala Lys Ile Ala Ala Leu Leu Glu Pro Glu Arg Ile Gly Val  
 1155 1160 1165  
 His Leu Ser Glu Glu Phe Gln Leu His Pro Glu Gln Ser Thr Asp Ala  
 1170 1175 1180  
 Ile Val Ile His His Pro Glu Ala Lys Tyr Phe Asn Ala Arg

1185 1190 10  
1195

<210> 3  
<211> 3537  
<212> DNA  
<213> Anabaena sp.

<220>  
<221> CDS  
<222> (1) .. (3534)  
<223> RAN03790

<400> 3

atg act cat cct ttc ctg aaa cgc ctg cac agt ccg gaa ctt ccg gtt 48  
Met Thr His Pro Phe Leu Lys Arg Leu His Ser Pro Glu Leu Pro Val  
1 5 10 15

atc gtc ttc gac ggt gca atg gga act aac cta caa acc caa aac ctc 96  
Ile Val Phe Asp Gly Ala Met Gly Thr Asn Leu Gln Thr Gln Asn Leu  
20 25 30

acg gct gag gat ttc ggc ggt gtg cag tat gaa ggt tgt aac gaa tac 144  
Thr Ala Glu Asp Phe Gly Gly Val Gln Tyr Glu Gly Cys Asn Glu Tyr  
35 40 45

cta gtc cac acc aaa ccc gaa gct gtc gcc aag gtt cac cgc gac ttt 192  
Leu Val His Thr Lys Pro Glu Ala Val Ala Lys Val His Arg Asp Phe  
50 55 60

ctc gct gtg ggt gca gat gtc atc gaa acc gac act ttc ggt gcg aca 240  
Leu Ala Val Gly Ala Asp Val Ile Glu Thr Asp Thr Phe Gly Ala Thr  
65 70 75 80

tcc att gtt ttg gcg gaa tat gac tta gca gac caa aca tat tac ctg 288  
Ser Ile Val Leu Ala Glu Tyr Asp Leu Ala Asp Gln Thr Tyr Tyr Leu  
85 90 95

aac aag aaa gcc gcc gaa ctg gcg aaa agt gtc gct gct gaa ttt tcc 336  
Asn Lys Lys Ala Ala Glu Leu Ala Lys Ser Val Ala Ala Glu Phe Ser  
100 105 110

aca cca gat aaa ccc cgg ttt gtt gct ggt tcc atc ggc ccc aca acc 384  
Thr Pro Asp Lys Pro Arg Phe Val Ala Gly Ser Ile Gly Pro Thr Thr  
115 120 125

aaa ctt ccc acc ttg gga cat atc gac ttt gac act ctc aaa act tgc 432  
Lys Leu Pro Thr Leu Gly His Ile Asp Phe Asp Thr Leu Lys Thr Cys  
130 135 140

ttt gct gaa caa gca gaa gcg ctg tta gat ggt ggc gtg gat tta ctt 480  
Phe Ala Glu Gln Ala Glu Ala Leu Leu Asp Gly Gly Val Asp Leu Leu  
145 150 155 160

ttg gtg gag act tgt caa gat gtg ctg caa atc aaa gcg gcg ctg aat 528  
Leu Val Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ala Ala Leu Asn  
165 170 175

ggg ata gaa gaa gtc ttt ggc aag aga ggg gaa cgc ata ccc ttg atg 576  
Gly Ile Glu Glu Val Phe Gly Lys Arg Gly Glu Arg Ile Pro Leu Met  
180 185 190



11

gtg tcc gtg aca atg gaa agc atg ggg aca atg ttg gtc ggt tcc gaa 624  
 Val Ser Val Thr Met Glu Ser Met Gly Thr Met Leu Val Gly Ser Glu  
 195 200 205

atc aac gcc gtc ctg aca att tta gaa cct ttc cca att gac att ctc 672  
 Ile Asn Ala Val Leu Thr Ile Leu Glu Pro Phe Pro Ile Asp Ile Leu  
 210 215 220

ggt ctg aac tgt gcc aca ggc cca gac ttg atg aaa cca cat att aaa 720  
 Gly Leu Asn Cys Ala Thr Gly Pro Asp Leu Met Lys Pro His Ile Lys  
 225 230 235 240

tat ttg gct gaa cat tcg ccg ttt gtg gtt tct tgt att cct aac gcg 768  
 Tyr Leu Ala Glu His Ser Pro Phe Val Val Ser Cys Ile Pro Asn Ala  
 245 250 255

ggt tta cca gaa aac gtt ggt ggt caa gca cat tat cgc tta aca cca 816  
 Gly Leu Pro Glu Asn Val Gly Gly Gln Ala His Tyr Arg Leu Thr Pro  
 260 265 270

atg gaa tta cgc atg gcg ttg atg cac ttt gtt gaa gat ttg ggt gtc 864  
 Met Glu Leu Arg Met Ala Leu Met His Phe Val Glu Asp Leu Gly Val  
 275 280 285

caa gtg atc ggg ggt tgc tgt ggg aca cgt cca gaa cac att caa caa 912  
 Gln Val Ile Gly Gly Cys Cys Gly Thr Arg Pro Glu His Ile Gln Gln  
 290 295 300

tta gca gaa att gcc aag gat tta aag cca aag gtg aga cag cca agt 960  
 Leu Ala Glu Ile Ala Lys Asp Leu Lys Pro Lys Val Arg Gln Pro Ser  
 305 310 315 320

tta gaa cct gcg gct gca tca ata tat agt act caa ccc tac gaa caa 1008  
 Leu Glu Pro Ala Ala Ala Ser Ile Tyr Ser Thr Gln Pro Tyr Glu Gln  
 325 330 335

gat aat tct ttc ttg att gtg ggt gaa cgc ctc aac gcc agt ggt tcc 1056  
 Asp Asn Ser Phe Leu Ile Val Gly Glu Arg Leu Asn Ala Ser Gly Ser  
 340 345 350

aag aaa tgc cgt gat ttg ctg aat gcg gaa gat tgg gac gga ttg gta 1104  
 Lys Lys Cys Arg Asp Leu Leu Asn Ala Glu Asp Trp Asp Gly Leu Val  
 355 360 365

tca atg gcg cga tcg caa gtc aag gaa ggc gca cat atc ctt gat gtc 1152  
 Ser Met Ala Arg Ser Gln Val Lys Glu Gly Ala His Ile Leu Asp Val  
 370 375 380

aac gtt gat tat gtg gga cgg gac ggt gtg cgg gat atg cac gaa cta 1200  
 Asn Val Asp Tyr Val Gly Arg Asp Gly Val Arg Asp Met His Glu Leu  
 385 390 395 400

gtt tcc cgc att gtg aat aat gtt aca ctc ccc tta atg ctc gac tcc 1248  
 Val Ser Arg Ile Val Asn Asn Val Thr Leu Pro Leu Met Leu Asp Ser  
 405 410 415

acc gaa tgg gaa aag atg gag gcg ggt tta aag gtg gct ggt ggt aag 1296  
 Thr Glu Trp Glu Lys Met Glu Ala Gly Leu Lys Val Ala Gly Gly Lys  
 420 425 430

tgt ttg ctg aac tcc acc aac tac gaa gat ggg gaa cca cgt ttc tta 1344  
 Cys Leu Leu Asn Ser Thr Asn Tyr Glu Asp Gly Glu Pro Arg Phe Leu

435	440	445	
aaa gtg ttg gag ttg gcg aag aaa tat ggc gcg ggt gtt gtt att ggc Lys Val Leu Glu Leu Ala Lys Lys Tyr Gly Ala Gly Val Val Ile Gly 450 455 460			1392
aca att gac gaa gaa ggg atg gcg cgg aca gcc gag aaa aag ttt caa Thr Ile Asp Glu Glu Gly Met Ala Arg Thr Ala Glu Lys Lys Phe Gln 465 470 475 480			1440
att gcc cag cgt gcc tat cgt caa tcg gta gaa tat ggg att ccc ccc Ile Ala Gln Arg Ala Tyr Arg Gln Ser Val Glu Tyr Gly Ile Pro Pro 485 490 495			1488
aca gaa ata ttc ttt gat acc tta gct tta cca att tct acc ggg att Thr Glu Ile Phe Phe Asp Thr Leu Ala Leu Pro Ile Ser Thr Gly Ile 500 505 510			1536
gaa gaa gac cgg gaa aat ggc aag gcg aca att gaa tca att agc cgt Glu Glu Asp Arg Glu Asn Gly Lys Ala Thr Ile Glu Ser Ile Ser Arg 515 520 525			1584
atc cgt aaa gaa ttg cca ggg tgt cat gtt att tta ggc gtg tca aat Ile Arg Lys Glu Leu Pro Gly Cys His Val Ile Leu Gly Val Ser Asn 530 535 540			1632
ata tcc ttt ggc tta aat tca gcc tcg cgg atg gtc tta aac tcc gtg Ile Ser Phe Gly Leu Asn Ser Ala Ser Arg Met Val Leu Asn Ser Val 545 550 555 560			1680
ttt ctc cat gaa gca atg act gct ggc atg gat gcg gcg atc gtc agt Phe Leu His Glu Ala Met Thr Ala Gly Met Asp Ala Ala Ile Val Ser 565 570 575			1728
gct agc aag att cta cca ctg tcg aag att gaa gag cgt cat caa gaa Ala Ser Lys Ile Leu Pro Leu Ser Lys Ile Glu Glu Arg His Gln Glu 580 585 590			1776
gtc tgc cgc cag tta att tat gac cag cgt aaa ttt gag ggt gat atc Val Cys Arg Gln Leu Ile Tyr Asp Gln Arg Lys Phe Glu Gly Asp Ile 595 600 605			1824
tgc atc tat gac ccc tta aca gaa cta act aaa ttg ttt gag gga gtc Cys Ile Tyr Asp Pro Leu Thr Glu Leu Thr Lys Leu Phe Glu Gly Val 610 615 620			1872
acc acc aaa cgt aac aaa ggc gtt gat gaa agc tta ccc atc gaa gaa Thr Thr Lys Arg Asn Lys Gly Val Asp Glu Ser Leu Pro Ile Glu Glu 625 630 635 640			1920
cga ctc aag cgt cac att atc gac ggc gaa cgc att ggt tta gaa gcg Arg Leu Lys Arg His Ile Ile Asp Gly Glu Arg Ile Gly Leu Glu Ala 645 650 655			1968
caa ctg aca aaa gcc tta gaa caa tat cca ccc cta gaa att atc aac Gln Leu Thr Lys Ala Leu Glu Gln Tyr Pro Pro Leu Glu Ile Ile Asn 660 665 670			2016
act ttc cta cta gat ggg atg aaa gta gtc ggg gaa ttg ttc ggt tca Thr Phe Leu Leu Asp Gly Met Lys Val Val Gly Glu Leu Phe Gly Ser 675 680 685			2064

13

gga caa atg cag cta cct ttc gtt tta cag tca gcc gaa acc atg aaa Gly Gln Met Gln Leu Pro Phe Val Leu Gln Ser Ala Glu Thr Met Lys 690 695 700	2112
gcg gcg gta gcc tac cta gaa ccg ttc atg gaa aaa tcg gaa agt ggc Ala Ala Val Ala Tyr Leu Glu Pro Phe Met Glu Lys Ser Glu Ser Gly 705 710 715 720	2160
aac aat gcc aaa ggt aaa gta att att gcc acc gtg aaa ggc gat gtt Asn Asn Ala Lys Gly Lys Val Ile Ile Ala Thr Val Lys Gly Asp Val 725 730 735	2208
cac gac att ggt aaa aac cta gta gac att atc ttg tcc aac aac ggc His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly 740 745 750	2256
tac aag gta att aac ctg gga att aaa cag ccg gtg gaa aat atc atc Tyr Lys Val Ile Asn Leu Gly Ile Lys Gln Pro Val Glu Asn Ile Ile 755 760 765	2304
gag gct tac aac caa cac aaa gct gat tgt att gcc atg agt ggc ttg Glu Ala Tyr Asn Gln His Lys Ala Asp Cys Ile Ala Met Ser Gly Leu 770 775 780	2352
ctg gta aaa tcc acc gca ttc atg aaa gaa aat ttg gag gtc ttc aac Leu Val Lys Ser Thr Ala Phe Met Lys Glu Asn Leu Glu Val Phe Asn 785 790 795 800	2400
gaa aaa ggc att aat gtt cct gta att tta ggt ggt gcg gca tta acc Glu Lys Gly Ile Asn Val Pro Val Ile Leu Gly Gly Ala Ala Leu Thr 805 810 815	2448
ccg aaa ttc gtg cat aaa gat tgc caa aat acc tac aaa ggt aaa gtc Pro Lys Phe Val His Lys Asp Cys Gln Asn Thr Tyr Lys Gly Lys Val 820 825 830	2496
att tat ggc aaa gat gct ttc tca gac ctg cat ttc atg gat aaa tta Ile Tyr Gly Lys Asp Ala Phe Ser Asp Leu His Phe Met Asp Lys Leu 835 840 845	2544
atg cca gcc aaa gcc act ggc aaa tgg gac aat tcc tta gga ttc ttg Met Pro Ala Lys Ala Thr Gly Lys Trp Asp Asn Ser Leu Gly Phe Leu 850 855 860	2592
gat gaa gta gaa acc gag gaa aca gaa cct acc aat cac aaa tcc cca Asp Glu Val Glu Thr Glu Glu Thr Glu Pro Thr Asn His Lys Ser Pro 865 870 875 880	2640
atc ccc agt ccc caa tcc cca gtc ccc agt ccc cag tcc cca gtc cct Ile Pro Ser Pro Gln Ser Pro Val Pro Ser Pro Gln Ser Pro Val Pro 885 890 895	2688
ata gac acc cga cgt tcc gaa gct gta gcc ata gac att ccc cgt ccc Ile Asp Thr Arg Arg Ser Glu Ala Val Ala Ile Asp Ile Pro Arg Pro 900 905 910	2736
aca cca cca ttc tgg gga acg caa tta tta cag cct agc gat att tcc Thr Pro Pro Phe Trp Gly Thr Gln Leu Leu Gln Pro Ser Asp Ile Ser 915 920 925	2784
tta gag gaa ata ttc tgg cac atg gat ttg caa gcc ttg att gcg gga	2832

14

Leu Glu Glu Ile Phe Trp His Met Asp Leu Gln Ala Leu Ile Ala Gly	
930 935 940	
caa tgg caa ttc cgc aaa ccc aaa gaa caa tca aag gaa gaa tat caa	2880
Gln Trp Gln Phe Arg Lys Pro Lys Glu Gln Ser Lys Glu Glu Tyr Gln	
945 950 955 960	
gct ttc ttg aat gag aaa gtg tat cca gtt cta gaa act tgg aaa cag	2928
Ala Phe Leu Asn Glu Lys Val Tyr Pro Val Leu Glu Thr Trp Lys Gln	
965 970 975	
cgc atc att gca gaa aac ttg tta cat ccc cag gta att tat ggg tat	2976
Arg Ile Ile Ala Glu Asn Leu Leu His Pro Gln Val Ile Tyr Gly Tyr	
980 985 990	
ttt cct tgt caa tct gag ggt aat act tta tat gtt tac gaa aca aac	3024
Phe Pro Cys Gln Ser Glu Gly Asn Thr Leu Tyr Val Tyr Glu Thr Asn	
995 1000 1005	
agc cca aat gcc aca gaa atc act cag ttt gaa ttc ccc cga caa aag	3072
Ser Pro Asn Ala Thr Glu Ile Thr Gln Phe Glu Phe Pro Arg Gln Lys	
1010 1015 1020	
tca tca aaa cga tta tgt att gcc gat ttc ttt gca ccg aaa gat tca	3120
Ser Ser Lys Arg Leu Cys Ile Ala Asp Phe Phe Ala Pro Lys Asp Ser	
1025 1030 1035 1040	
gga atc att gat gtc ttc ccc atg cag gcg gtg act gta ggc gaa att	3168
Gly Ile Ile Asp Val Phe Pro Met Gln Ala Val Thr Val Gly Glu Ile	
1045 1050 1055	
gct aca gag ttc gcg caa aaa ttg ttt gca aac aat caa tac act gat	3216
Ala Thr Glu Phe Ala Gln Lys Leu Phe Ala Asn Asn Gln Tyr Thr Asp	
1060 1065 1070	
tat ctg tat ttt cac ggt ttg gcg gtg caa gta gca gaa gcc ttg gcc	3264
Tyr Leu Tyr Phe His Gly Leu Ala Val Gln Val Ala Glu Ala Leu Ala	
1075 1080 1085	
gag tgg aca cac gcc aga atc cgc cgt gag tta ggg ttc ggt gct gaa	3312
Glu Trp Thr His Ala Arg Ile Arg Arg Glu Leu Gly Phe Gly Ala Glu	
1090 1095 1100	
gaa ccg gat aat atc cgg gat att ttg gca caa cgc tat cag ggt tcc	3360
Glu Pro Asp Asn Ile Arg Asp Ile Leu Ala Gln Arg Tyr Gln Gly Ser	
1105 1110 1115 1120	
cgg tat agt ttt ggc tac cca gct tgt ccc aat att caa gac cag ttt	3408
Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn Ile Gln Asp Gln Phe	
1125 1130 1135	
aag cag ctg gat ttg ttg gag act agc aga att aac tta tac atg gat	3456
Lys Gln Leu Asp Leu Leu Glu Thr Ser Arg Ile Asn Leu Tyr Met Asp	
1140 1145 1150	
gaa agt gag caa ctt tat cca gaa cag tct acg acg gcg att att act	3504
Glu Ser Glu Gln Leu Tyr Pro Glu Gln Ser Thr Thr Ala Ile Ile Thr	
1155 1160 1165	
tat cac cca gta gct aag tac ttc acc gcg taa	3537
Tyr His Pro Val Ala Lys Tyr Phe Thr Ala	
1170 1175	

<210> 4  
 <211> 1178  
 <212> PRT  
 <213> Anabaena sp.

<400> 4

```

Met Thr His Pro Phe Leu Lys Arg Leu His Ser Pro Glu Leu Pro Val
 1           5           10           15

Ile Val Phe Asp Gly Ala Met Gly Thr Asn Leu Gln Thr Gln Asn Leu
      20           25           30

Thr Ala Glu Asp Phe Gly Gly Val Gln Tyr Glu Gly Cys Asn Glu Tyr
      35           40           45

Leu Val His Thr Lys Pro Glu Ala Val Ala Lys Val His Arg Asp Phe
      50           55           60

Leu Ala Val Gly Ala Asp Val Ile Glu Thr Asp Thr Phe Gly Ala Thr
      65           70           75           80

Ser Ile Val Leu Ala Glu Tyr Asp Leu Ala Asp Gln Thr Tyr Tyr Leu
      85           90           95

Asn Lys Lys Ala Ala Glu Leu Ala Lys Ser Val Ala Ala Glu Phe Ser
      100           105           110

Thr Pro Asp Lys Pro Arg Phe Val Ala Gly Ser Ile Gly Pro Thr Thr
      115           120           125

Lys Leu Pro Thr Leu Gly His Ile Asp Phe Asp Thr Leu Lys Thr Cys
      130           135           140

Phe Ala Glu Gln Ala Glu Ala Leu Leu Asp Gly Gly Val Asp Leu Leu
      145           150           155           160

Leu Val Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ala Ala Leu Asn
      165           170           175

Gly Ile Glu Glu Val Phe Gly Lys Arg Gly Glu Arg Ile Pro Leu Met
      180           185           190

Val Ser Val Thr Met Glu Ser Met Gly Thr Met Leu Val Gly Ser Glu
      195           200           205

Ile Asn Ala Val Leu Thr Ile Leu Glu Pro Phe Pro Ile Asp Ile Leu
      210           215           220

Gly Leu Asn Cys Ala Thr Gly Pro Asp Leu Met Lys Pro His Ile Lys
      225           230           235           240

Tyr Leu Ala Glu His Ser Pro Phe Val Val Ser Cys Ile Pro Asn Ala
      245           250           255

Gly Leu Pro Glu Asn Val Gly Gly Gln Ala His Tyr Arg Leu Thr Pro
      260           265           270

Met Glu Leu Arg Met Ala Leu Met His Phe Val Glu Asp Leu Gly Val
      275           280           285
  
```

Gln Val Ile Gly Gly Cys Cys Gly Thr Arg Pro Glu His Ile Gln Gln  
 290 295 300  
 Leu Ala Glu Ile Ala Lys Asp Leu Lys Pro Lys Val Arg Gln Pro Ser  
 305 310 315 320  
 Leu Glu Pro Ala Ala Ala Ser Ile Tyr Ser Thr Gln Pro Tyr Glu Gln  
 325 330 335  
 Asp Asn Ser Phe Leu Ile Val Gly Glu Arg Leu Asn Ala Ser Gly Ser  
 340 345 350  
 Lys Lys Cys Arg Asp Leu Leu Asn Ala Glu Asp Trp Asp Gly Leu Val  
 355 360 365  
 Ser Met Ala Arg Ser Gln Val Lys Glu Gly Ala His Ile Leu Asp Val  
 370 375 380  
 Asn Val Asp Tyr Val Gly Arg Asp Gly Val Arg Asp Met His Glu Leu  
 385 390 395 400  
 Val Ser Arg Ile Val Asn Asn Val Thr Leu Pro Leu Met Leu Asp Ser  
 405 410 415  
 Thr Glu Trp Glu Lys Met Glu Ala Gly Leu Lys Val Ala Gly Gly Lys  
 420 425 430  
 Cys Leu Leu Asn Ser Thr Asn Tyr Glu Asp Gly Glu Pro Arg Phe Leu  
 435 440 445  
 Lys Val Leu Glu Leu Ala Lys Lys Tyr Gly Ala Gly Val Val Ile Gly  
 450 455 460  
 Thr Ile Asp Glu Glu Gly Met Ala Arg Thr Ala Glu Lys Lys Phe Gln  
 465 470 475 480  
 Ile Ala Gln Arg Ala Tyr Arg Gln Ser Val Glu Tyr Gly Ile Pro Pro  
 485 490 495  
 Thr Glu Ile Phe Phe Asp Thr Leu Ala Leu Pro Ile Ser Thr Gly Ile  
 500 505 510  
 Glu Glu Asp Arg Glu Asn Gly Lys Ala Thr Ile Glu Ser Ile Ser Arg  
 515 520 525  
 Ile Arg Lys Glu Leu Pro Gly Cys His Val Ile Leu Gly Val Ser Asn  
 530 535 540  
 Ile Ser Phe Gly Leu Asn Ser Ala Ser Arg Met Val Leu Asn Ser Val  
 545 550 555 560  
 Phe Leu His Glu Ala Met Thr Ala Gly Met Asp Ala Ala Ile Val Ser  
 565 570 575  
 Ala Ser Lys Ile Leu Pro Leu Ser Lys Ile Glu Glu Arg His Gln Glu  
 580 585 590  
 Val Cys Arg Gln Leu Ile Tyr Asp Gln Arg Lys Phe Glu Gly Asp Ile  
 595 600 605  
 Cys Ile Tyr Asp Pro Leu Thr Glu Leu Thr Lys Leu Phe Glu Gly Val

610                      615                      620  
 Thr Thr Lys Arg Asn Lys Gly Val Asp Glu Ser Leu Pro Ile Glu Glu  
 625                      630                      635                      640  
 Arg Leu Lys Arg His Ile Ile Asp Gly Glu Arg Ile Gly Leu Glu Ala  
                     645                      650                      655  
 Gln Leu Thr Lys Ala Leu Glu Gln Tyr Pro Pro Leu Glu Ile Ile Asn  
                     660                      665                      670  
 Thr Phe Leu Leu Asp Gly Met Lys Val Val Gly Glu Leu Phe Gly Ser  
                     675                      680                      685  
 Gly Gln Met Gln Leu Pro Phe Val Leu Gln Ser Ala Glu Thr Met Lys  
                     690                      695                      700  
 Ala Ala Val Ala Tyr Leu Glu Pro Phe Met Glu Lys Ser Glu Ser Gly  
 705                      710                      715                      720  
 Asn Asn Ala Lys Gly Lys Val Ile Ile Ala Thr Val Lys Gly Asp Val  
                     725                      730                      735  
 His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly  
                     740                      745                      750  
 Tyr Lys Val Ile Asn Leu Gly Ile Lys Gln Pro Val Glu Asn Ile Ile  
                     755                      760                      765  
 Glu Ala Tyr Asn Gln His Lys Ala Asp Cys Ile Ala Met Ser Gly Leu  
                     770                      775                      780  
 Leu Val Lys Ser Thr Ala Phe Met Lys Glu Asn Leu Glu Val Phe Asn  
 785                      790                      795                      800  
 Glu Lys Gly Ile Asn Val Pro Val Ile Leu Gly Gly Ala Ala Leu Thr  
                     805                      810                      815  
 Pro Lys Phe Val His Lys Asp Cys Gln Asn Thr Tyr Lys Gly Lys Val  
                     820                      825                      830  
 Ile Tyr Gly Lys Asp Ala Phe Ser Asp Leu His Phe Met Asp Lys Leu  
                     835                      840                      845  
 Met Pro Ala Lys Ala Thr Gly Lys Trp Asp Asn Ser Leu Gly Phe Leu  
                     850                      855                      860  
 Asp Glu Val Glu Thr Glu Glu Thr Glu Pro Thr Asn His Lys Ser Pro  
 865                      870                      875                      880  
 Ile Pro Ser Pro Gln Ser Pro Val Pro Ser Pro Gln Ser Pro Val Pro  
                     885                      890                      895  
 Ile Asp Thr Arg Arg Ser Glu Ala Val Ala Ile Asp Ile Pro Arg Pro  
                     900                      905                      910  
 Thr Pro Pro Phe Trp Gly Thr Gln Leu Leu Gln Pro Ser Asp Ile Ser  
                     915                      920                      925  
 Leu Glu Glu Ile Phe Trp His Met Asp Leu Gln Ala Leu Ile Ala Gly  
                     930                      935                      940

18

Gln Trp Gln Phe Arg Lys Pro Lys Glu Gln Ser Lys Glu Glu Tyr Gln  
 945 950 955 960  
 Ala Phe Leu Asn Glu Lys Val Tyr Pro Val Leu Glu Thr Trp Lys Gln  
 965 970 975  
 Arg Ile Ile Ala Glu Asn Leu Leu His Pro Gln Val Ile Tyr Gly Tyr  
 980 985 990  
 Phe Pro Cys Gln Ser Glu Gly Asn Thr Leu Tyr Val Tyr Glu Thr Asn  
 995 1000 1005  
 Ser Pro Asn Ala Thr Glu Ile Thr Gln Phe Glu Phe Pro Arg Gln Lys  
 1010 1015 1020  
 Ser Ser Lys Arg Leu Cys Ile Ala Asp Phe Phe Ala Pro Lys Asp Ser  
 1025 1030 1035 1040  
 Gly Ile Ile Asp Val Phe Pro Met Gln Ala Val Thr Val Gly Glu Ile  
 1045 1050 1055  
 Ala Thr Glu Phe Ala Gln Lys Leu Phe Ala Asn Asn Gln Tyr Thr Asp  
 1060 1065 1070  
 Tyr Leu Tyr Phe His Gly Leu Ala Val Gln Val Ala Glu Ala Leu Ala  
 1075 1080 1085  
 Glu Trp Thr His Ala Arg Ile Arg Arg Glu Leu Gly Phe Gly Ala Glu  
 1090 1095 1100  
 Glu Pro Asp Asn Ile Arg Asp Ile Leu Ala Gln Arg Tyr Gln Gly Ser  
 1105 1110 1115 1120  
 Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn Ile Gln Asp Gln Phe  
 1125 1130 1135  
 Lys Gln Leu Asp Leu Leu Glu Thr Ser Arg Ile Asn Leu Tyr Met Asp  
 1140 1145 1150  
 Glu Ser Glu Gln Leu Tyr Pro Glu Gln Ser Thr Thr Ala Ile Ile Thr  
 1155 1160 1165  
 Tyr His Pro Val Ala Lys Tyr Phe Thr Ala  
 1170 1175

<210> 5  
 <211> 3588  
 <212> DNA  
 <213> Synechocystis sp.

<220>  
 <221> CDS  
 <222> (1) .. (3585)  
 <223> RCY35965

<400> 5  
 atg aaa agt gct ttt tta gac cgt atc cac agt ccc gat cgc ccg gta 48  
 Met Lys Ser Ala Phe Leu Asp Arg Ile His Ser Pro Asp Arg Pro Val  
 1 5 10 15  
 tta gtc ttt gac ggg gct atg ggt aca aac ctg cag gta cag aac cta 96



19

Leu Val Phe Asp Gly Ala Met Gly Thr Asn Leu Gln Val Gln Asn Leu	
20 25 30	
acg gcg gcg gat ttt ggt ggg gcg gaa tac gaa ggt tgc aat gaa tat	144
Thr Ala Ala Asp Phe Gly Gly Ala Glu Tyr Glu Gly Cys Asn Glu Tyr	
35 40 45	
tta gtc cat acc aag cca gag gcc gtg gct acg gtg cat cgt gct ttt	192
Leu Val His Thr Lys Pro Glu Ala Val Ala Thr Val His Arg Ala Phe	
50 55 60	
tac gaa gcg ggg gcc gat gtc gtg gaa acg gat act ttt ggg gga acg	240
Tyr Glu Ala Gly Ala Asp Val Val Glu Thr Asp Thr Phe Gly Gly Thr	
65 70 75 80	
ccc ctg gtg ctg gcg gag tac gat tta gca gac caa agt tat tac tta	288
Pro Leu Val Leu Ala Glu Tyr Asp Leu Ala Asp Gln Ser Tyr Tyr Leu	
85 90 95	
aat aaa gca gcg gcg gag ttg gcc aag gcg gta gca gcg gaa ttt tct	336
Asn Lys Ala Ala Ala Glu Leu Ala Lys Ala Val Ala Ala Glu Phe Ser	
100 105 110	
acc cca gaa aag cct cga ttc gtg gcc ggc tcc atg gga cca ggc acc	384
Thr Pro Glu Lys Pro Arg Phe Val Ala Gly Ser Met Gly Pro Gly Thr	
115 120 125	
aag cta ccc acc cta ggt cat gtg gac tac gat agt ctc aag gat gcc	432
Lys Leu Pro Thr Leu Gly His Val Asp Tyr Asp Ser Leu Lys Asp Ala	
130 135 140	
tat gtg gtt cag gtg cgg ggt tta tac gat ggc gga gtg gat tta ttg	480
Tyr Val Val Gln Val Arg Gly Leu Tyr Asp Gly Gly Val Asp Leu Leu	
145 150 155 160	
cta gtg gaa acc tgc cag gat gtg ctg caa att aaa gcg gcc ttg aac	528
Leu Val Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ala Ala Leu Asn	
165 170 175	
gcc att gaa cag gtc ttt gcc gaa aaa ggc gat cgc cta ccg ttg atg	576
Ala Ile Glu Gln Val Phe Ala Glu Lys Gly Asp Arg Leu Pro Leu Met	
180 185 190	
gtg tca gta acc atg gaa acc atg ggg acc atg ctg gtg ggt acg gag	624
Val Ser Val Thr Met Glu Thr Met Gly Thr Met Leu Val Gly Thr Glu	
195 200 205	
atg gcg gcg gcc ctg gcc att ttg gag ccc tat ccc atc gat att ttg	672
Met Ala Ala Ala Leu Ala Ile Leu Glu Pro Tyr Pro Ile Asp Ile Leu	
210 215 220	
ggg cta aac tgc gcc acc ggg cca gat ttg atg aag gaa cac gtt aaa	720
Gly Leu Asn Cys Ala Thr Gly Pro Asp Leu Met Lys Glu His Val Lys	
225 230 235 240	
tat ctt tcc gaa cat tcc ccc ttt gtg gtg tcc tgt att ccc aat gct	768
Tyr Leu Ser Glu His Ser Pro Phe Val Val Ser Cys Ile Pro Asn Ala	
245 250 255	
ggt ttg cca gaa aac gtt ggc ggt caa gct ttt tat cgc ctc acc ccg	816
Gly Leu Pro Glu Asn Val Gly Gly Gln Ala Phe Tyr Arg Leu Thr Pro	
260 265 270	

## 20

atg gaa ctg caa atg tcc ctg atg cac ttc atc gaa gac ctg gga gta Met Glu Leu Gln Met Ser Leu Met His Phe Ile Glu Asp Leu Gly Val 275 280 285	864
cag gta att ggt ggt tgt tgt ggc act aga ccc gat cac atc aag gcc Gln Val Ile Gly Gly Cys Cys Gly Thr Arg Pro Asp His Ile Lys Ala 290 295 300	912
ctg gcg gat att gcc aag gat ctc cag ccc aaa caa cgc caa cct cac Leu Ala Asp Ile Ala Lys Asp Leu Gln Pro Lys Gln Arg Gln Pro His 305 310 315 320	960
tac gaa ccc agc gcc gct tcc att tat tcc acc caa acc tac gcc caa Tyr Glu Pro Ser Ala Ala Ser Ile Tyr Ser Thr Gln Thr Tyr Ala Gln 325 330 335	1008
gaa aat tct ttt tta atc att ggc gaa cgg ctc aat gcc agt ggc tcg Glu Asn Ser Phe Leu Ile Ile Gly Glu Arg Leu Asn Ala Ser Gly Ser 340 345 350	1056
aaa aaa tgt cga gat ctg ctc aat gct gaa gat tgg gac agc cta gtt Lys Lys Cys Arg Asp Leu Leu Asn Ala Glu Asp Trp Asp Ser Leu Val 355 360 365	1104
tcc ctg gct aaa tcc caa gtc aag gaa gga gcc caa atc ctt gac gtc Ser Leu Ala Lys Ser Gln Val Lys Glu Gly Ala Gln Ile Leu Asp Val 370 375 380	1152
aac gtg gat tac gtt ggt cga gat ggg gta agg gac atg aaa gaa tta Asn Val Asp Tyr Val Gly Arg Asp Gly Val Arg Asp Met Lys Glu Leu 385 390 395 400	1200
gct tcc cga cta gtc aat aat gtc acc ctg ccg ttg atg ttg gac tcc Ala Ser Arg Leu Val Asn Asn Val Thr Leu Pro Leu Met Leu Asp Ser 405 410 415	1248
acc gaa tgg caa aaa atg gag gcg ggt tta aaa gtt gca ggg gga aaa Thr Glu Trp Gln Lys Met Glu Ala Gly Leu Lys Val Ala Gly Gly Lys 420 425 430	1296
tgt att ctc aat tcc acc aac tac gaa gac ggg gaa gaa cgg ttt tat Cys Ile Leu Asn Ser Thr Asn Tyr Glu Asp Gly Glu Glu Arg Phe Tyr 435 440 445	1344
aaa gtg tta gaa att gcc aaa gaa tat gga gct ggt att gtc att ggc Lys Val Leu Glu Ile Ala Lys Glu Tyr Gly Ala Gly Ile Val Ile Gly 450 455 460	1392
acc atc gat gaa gat ggc atg gga cgc act gca gat aaa aaa ttt gag Thr Ile Asp Glu Asp Gly Met Gly Arg Thr Ala Asp Lys Lys Phe Glu 465 470 475 480	1440
att gcc aaa cgg gcc tac gaa gcg gcg atc gcc ttt ggc att ccg gcc Ile Ala Lys Arg Ala Tyr Glu Ala Ala Ile Ala Phe Gly Ile Pro Ala 485 490 495	1488
aca gaa att ttc ttt gat cct tta gct ctg cct att tcc acc ggc att Thr Glu Ile Phe Phe Asp Pro Leu Ala Leu Pro Ile Ser Thr Gly Ile 500 505 510	1536
gaa gaa gac agg gag aac ggt aaa gcc acc gtg gat gct atc cgc aga	1584

## 21

Glu Glu Asp Arg Glu Asn Gly Lys Ala Thr Val Asp Ala Ile Arg Arg 515 520 525	
att cgc cag gaa ttg ccc gat tgt cat att ttg ttg ggg gtt tct aac Ile Arg Gln Glu Leu Pro Asp Cys His Ile Leu Leu Gly Val Ser Asn 530 535 540	1632
gtt tcc ttt ggc ttg aat ccc gcc gct cgc cag gta ctc aat tcc atc Val Ser Phe Gly Leu Asn Pro Ala Ala Arg Gln Val Leu Asn Ser Ile 545 550 555 560	1680
ttt ctc cac gaa tgt atg cag gtg ggc atg gat gcg gcc att gtc agt Phe Leu His Glu Cys Met Gln Val Gly Met Asp Ala Ala Ile Val Ser 565 570 575	1728
gcc aat aag att tta ccc ctg gca aaa att gac cca gaa caa caa caa Ala Asn Lys Ile Leu Pro Leu Ala Lys Ile Asp Pro Glu Gln Gln Gln 580 585 590	1776
gtc tgt cta gat tta atc tat gac cgc cgg gaa ttt gaa gga gag cgc Val Cys Leu Asp Leu Ile Tyr Asp Arg Arg Glu Phe Glu Gly Glu Arg 595 600 605	1824
tgt aca tat gac ccg tta acc aaa ctc acc act tta ttt gaa ggt aaa Cys Thr Tyr Asp Pro Leu Thr Lys Leu Thr Thr Leu Phe Glu Gly Lys 610 615 620	1872
acc acc aaa cgg gat aaa tcc ggt gat gcc aat tta ccg gtg gaa gaa Thr Thr Lys Arg Asp Lys Ser Gly Asp Ala Asn Leu Pro Val Glu Glu 625 630 635 640	1920
aga tta aaa cgc cac atc att gat ggg gaa aga ttg ggc tta gaa gag Arg Leu Lys Arg His Ile Ile Asp Gly Glu Arg Leu Gly Leu Glu Glu 645 650 655	1968
gcc ctc aat gaa gct tta aaa ctt tac gct ccc tta gat atc att aac Ala Leu Asn Glu Ala Leu Lys Leu Tyr Ala Pro Leu Asp Ile Ile Asn 660 665 670	2016
atc tat ttg ttg gat ggc atg aaa gtg gtg ggg gaa cta ttt ggt tcc Ile Tyr Leu Leu Asp Gly Met Lys Val Val Gly Glu Leu Phe Gly Ser 675 680 685	2064
ggg caa atg cag ttg ccc ttt gtg ttg cag tcg gcc caa acc atg aaa Gly Gln Met Gln Leu Pro Phe Val Leu Gln Ser Ala Gln Thr Met Lys 690 695 700	2112
gcg gcg gtg gct ttt tta gaa ccc cat atg gat aag gat gat tcc gcc Ala Ala Val Ala Phe Leu Glu Pro His Met Asp Lys Asp Asp Ser Ala 705 710 715 720	2160
gac aat gct aag ggt act ttt tta att gcc act gtt aag ggg gat gtc Asp Asn Ala Lys Gly Thr Phe Leu Ile Ala Thr Val Lys Gly Asp Val 725 730 735	2208
cat gat att ggc aaa aac tta gtg gat att atc ctt tcc aac aat ggc His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly 740 745 750	2256
tat cga gtg gtc aac cta ggc att aaa cag cca gtg gaa aat att atc Tyr Arg Val Val Asn Leu Gly Ile Lys Gln Pro Val Glu Asn Ile Ile 755 760 765	2304

## 22

gaa gcc tac aaa aaa cac agg ccc gat tgc att gcc atg agt ggt ttg Glu Ala Tyr Lys Lys His Arg Pro Asp Cys Ile Ala Met Ser Gly Leu 770 775 780	2352
ttg gtc aaa tca act gct ttt atg aag gaa aat tta gaa gtt ttc aac Leu Val Lys Ser Thr Ala Phe Met Lys Glu Asn Leu Glu Val Phe Asn 785 790 795 800	2400
caa gag ggc att act gtt ccc gtc att ctt ggt ggt gct gct tta acg Gln Glu Gly Ile Thr Val Pro Val Ile Leu Gly Gly Ala Ala Leu Thr 805 810 815	2448
cct aaa ttt gtt cac cag gac tgc caa aat acc tac aaa ggc caa gta Pro Lys Phe Val His Gln Asp Cys Gln Asn Thr Tyr Lys Gly Gln Val 820 825 830	2496
att tac ggc aaa gat gcg ttc gcc gat tta cat ttc atg gat aag cta Ile Tyr Gly Lys Asp Ala Phe Ala Asp Leu His Phe Met Asp Lys Leu 835 840 845	2544
atg ccc gct aaa aat agc cac aat tgg gat gat ttc cag ggc ttt tta Met Pro Ala Lys Asn Ser His Asn Trp Asp Asp Phe Gln Gly Phe Leu 850 855 860	2592
ggg gaa tat gca acg gaa aat ggc cat aat gtg acc act gat gat ggt Gly Glu Tyr Ala Thr Glu Asn Gly His Asn Val Thr Thr Asp Asp Gly 865 870 875 880	2640
gct aaa act aat ttt ggc att gaa gaa gaa aaa tta att gac gct agt Ala Lys Thr Asn Phe Gly Ile Glu Glu Glu Lys Leu Ile Asp Ala Ser 885 890 895	2688
gag cag tct agg gag ccg gag gta att gat act gtt cgt tct gaa gcg Glu Gln Ser Arg Glu Pro Glu Val Ile Asp Thr Val Arg Ser Glu Ala 900 905 910	2736
gtg gac cct gat cta gaa aga cct gtg cca cct ttt tgg ggc act aaa Val Asp Pro Asp Leu Glu Arg Pro Val Pro Pro Phe Trp Gly Thr Lys 915 920 925	2784
att ttg caa tcc agt gat att tcc ctc gat gaa gtc ttc cct tta ctg Ile Leu Gln Ser Ser Asp Ile Ser Leu Asp Glu Val Phe Pro Leu Leu 930 935 940	2832
gat tta caa gca tta ttt gtt ggt cag tgg cag ttt cgc aaa cct agg Asp Leu Gln Ala Leu Phe Val Gly Gln Trp Gln Phe Arg Lys Pro Arg 945 950 955 960	2880
gag caa tcc agg gaa gaa tac gag caa ttc cta gcg gaa aaa gtt cat Glu Gln Ser Arg Glu Glu Tyr Glu Gln Phe Leu Ala Glu Lys Val His 965 970 975	2928
ccc att ttg gct gag tgg aaa ggt aag gtc atg gca gaa aat tta ctc Pro Ile Leu Ala Glu Trp Lys Gly Lys Val Met Ala Glu Asn Leu Leu 980 985 990	2976
cat cct acg gtg gtt tat ggt tat ttt ccc tgt caa tcc cag ggc aat His Pro Thr Val Val Tyr Gly Tyr Phe Pro Cys Gln Ser Gln Gly Asn 995 1000 1005	3024
acc ttg tta att tat gac cca gaa ttg gtc agc caa aat aat ggc caa	3072

## 23

Thr Leu Leu Ile Tyr Asp Pro Glu Leu Val Ser Gln Asn Asn Gly Gln  
 1010 1015 1020  
 att ccc cca gac gca acg gcg atc gcc aaa ttt gag ttt ccc cgg caa 3120  
 Ile Pro Pro Asp Ala Thr Ala Ile Ala Lys Phe Glu Phe Pro Arg Gln  
 1025 1030 1035 1040  
 aaa tca ggg cgg cgg ctc tgt att gcg gac ttt ttt gct tca aaa gaa 3168  
 Lys Ser Gly Arg Arg Leu Cys Ile Ala Asp Phe Phe Ala Ser Lys Glu  
 1045 1050 1055  
 tcg ggg att act gat gtt ttt cct ttg caa gcg gtt aca gtg ggg gaa 3216  
 Ser Gly Ile Thr Asp Val Phe Pro Leu Gln Ala Val Thr Val Gly Glu  
 1060 1065 1070  
 atc gcg acg gaa tat gca agg aaa ctt ttt gct ggc gat aat tac acc 3264  
 Ile Ala Thr Glu Tyr Ala Arg Lys Leu Phe Ala Gly Asp Asn Tyr Thr  
 1075 1080 1085  
 gat tac ctc tac ttc cac ggc atg gcg gtg cag atg gcg gaa gct tta 3312  
 Asp Tyr Leu Tyr Phe His Gly Met Ala Val Gln Met Ala Glu Ala Leu  
 1090 1095 1100  
 gcg gag tgg act cac caa cgg ata cgt cag gaa ttg ggc ttt ggc cat 3360  
 Ala Glu Trp Thr His Gln Arg Ile Arg Gln Glu Leu Gly Phe Gly His  
 1105 1110 1115 1120  
 tta gat cca gat aac atc cgt gat ctt ctc cag caa cgt tac caa ggt 3408  
 Leu Asp Pro Asp Asn Ile Arg Asp Leu Leu Gln Gln Arg Tyr Gln Gly  
 1125 1130 1135  
 tcc cgc tac agt ttt ggt tat ccc gct tgt ccc aac atg cag gat caa 3456  
 Ser Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn Met Gln Asp Gln  
 1140 1145 1150  
 tac aca caa tta gaa ttg tta caa acc gaa cga att ggc ttg tat atg 3504  
 Tyr Thr Gln Leu Glu Leu Leu Thr Glu Arg Ile Gly Leu Tyr Met  
 1155 1160 1165  
 gat gaa agt gaa cag gtt tat cca gaa caa tcc acc acg gcg att att 3552  
 Asp Glu Ser Glu Gln Val Tyr Pro Glu Gln Ser Thr Thr Ala Ile Ile  
 1170 1175 1180  
 tcc tat cat cct gcg gct aaa tat ttc agc gct taa 3588  
 Ser Tyr His Pro Ala Ala Lys Tyr Phe Ser Ala  
 1185 1190 1195  
 <210> 6  
 <211> 1195  
 <212> PRT  
 <213> Synechocystis sp.  
 <400> 6  
 Met Lys Ser Ala Phe Leu Asp Arg Ile His Ser Pro Asp Arg Pro Val  
 1 5 10 15  
 Leu Val Phe Asp Gly Ala Met Gly Thr Asn Leu Gln Val Gln Asn Leu  
 20 25 30  
 Thr Ala Ala Asp Phe Gly Gly Ala Glu Tyr Glu Gly Cys Asn Glu Tyr

24

35                      40                      45  
 Leu Val His Thr Lys Pro Glu Ala Val Ala Thr Val His Arg Ala Phe  
   50                      55                      60  
 Tyr Glu Ala Gly Ala Asp Val Val Glu Thr Asp Thr Phe Gly Gly Thr  
   65                      70                      75                      80  
 Pro Leu Val Leu Ala Glu Tyr Asp Leu Ala Asp Gln Ser Tyr Tyr Leu  
                     85                      90                      95  
 Asn Lys Ala Ala Ala Glu Leu Ala Lys Ala Val Ala Ala Glu Phe Ser  
                     100                      105                      110  
 Thr Pro Glu Lys Pro Arg Phe Val Ala Gly Ser Met Gly Pro Gly Thr  
                     115                      120                      125  
 Lys Leu Pro Thr Leu Gly His Val Asp Tyr Asp Ser Leu Lys Asp Ala  
                     130                      135                      140  
 Tyr Val Val Gln Val Arg Gly Leu Tyr Asp Gly Gly Val Asp Leu Leu  
                     145                      150                      155                      160  
 Leu Val Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ala Ala Leu Asn  
                     165                      170                      175  
 Ala Ile Glu Gln Val Phe Ala Glu Lys Gly Asp Arg Leu Pro Leu Met  
                     180                      185                      190  
 Val Ser Val Thr Met Glu Thr Met Gly Thr Met Leu Val Gly Thr Glu  
                     195                      200                      205  
 Met Ala Ala Ala Leu Ala Ile Leu Glu Pro Tyr Pro Ile Asp Ile Leu  
                     210                      215                      220  
 Gly Leu Asn Cys Ala Thr Gly Pro Asp Leu Met Lys Glu His Val Lys  
                     225                      230                      235                      240  
 Tyr Leu Ser Glu His Ser Pro Phe Val Val Ser Cys Ile Pro Asn Ala  
                     245                      250                      255  
 Gly Leu Pro Glu Asn Val Gly Gly Gln Ala Phe Tyr Arg Leu Thr Pro  
                     260                      265                      270  
 Met Glu Leu Gln Met Ser Leu Met His Phe Ile Glu Asp Leu Gly Val  
                     275                      280                      285  
 Gln Val Ile Gly Gly Cys Cys Gly Thr Arg Pro Asp His Ile Lys Ala  
                     290                      295                      300  
 Leu Ala Asp Ile Ala Lys Asp Leu Gln Pro Lys Gln Arg Gln Pro His  
                     305                      310                      315                      320  
 Tyr Glu Pro Ser Ala Ala Ser Ile Tyr Ser Thr Gln Thr Tyr Ala Gln  
                     325                      330                      335  
 Glu Asn Ser Phe Leu Ile Ile Gly Glu Arg Leu Asn Ala Ser Gly Ser  
                     340                      345                      350  
 Lys Lys Cys Arg Asp Leu Leu Asn Ala Glu Asp Trp Asp Ser Leu Val  
                     355                      360                      365

25

Ser Leu Ala Lys Ser Gln Val Lys Glu Gly Ala Gln Ile Leu Asp Val  
 370 375 380  
 Asn Val Asp Tyr Val Gly Arg Asp Gly Val Arg Asp Met Lys Glu Leu  
 385 390 395 400  
 Ala Ser Arg Leu Val Asn Asn Val Thr Leu Pro Leu Met Leu Asp Ser  
 405 410 415  
 Thr Glu Trp Gln Lys Met Glu Ala Gly Leu Lys Val Ala Gly Gly Lys  
 420 425 430  
 Cys Ile Leu Asn Ser Thr Asn Tyr Glu Asp Gly Glu Glu Arg Phe Tyr  
 435 440 445  
 Lys Val Leu Glu Ile Ala Lys Glu Tyr Gly Ala Gly Ile Val Ile Gly  
 450 455 460  
 Thr Ile Asp Glu Asp Gly Met Gly Arg Thr Ala Asp Lys Lys Phe Glu  
 465 470 475 480  
 Ile Ala Lys Arg Ala Tyr Glu Ala Ala Ile Ala Phe Gly Ile Pro Ala  
 485 490 495  
 Thr Glu Ile Phe Phe Asp Pro Leu Ala Leu Pro Ile Ser Thr Gly Ile  
 500 505 510  
 Glu Glu Asp Arg Glu Asn Gly Lys Ala Thr Val Asp Ala Ile Arg Arg  
 515 520 525  
 Ile Arg Gln Glu Leu Pro Asp Cys His Ile Leu Leu Gly Val Ser Asn  
 530 535 540  
 Val Ser Phe Gly Leu Asn Pro Ala Ala Arg Gln Val Leu Asn Ser Ile  
 545 550 555 560  
 Phe Leu His Glu Cys Met Gln Val Gly Met Asp Ala Ala Ile Val Ser  
 565 570 575  
 Ala Asn Lys Ile Leu Pro Leu Ala Lys Ile Asp Pro Glu Gln Gln Gln  
 580 585 590  
 Val Cys Leu Asp Leu Ile Tyr Asp Arg Arg Glu Phe Glu Gly Glu Arg  
 595 600 605  
 Cys Thr Tyr Asp Pro Leu Thr Lys Leu Thr Thr Leu Phe Glu Gly Lys  
 610 615 620  
 Thr Thr Lys Arg Asp Lys Ser Gly Asp Ala Asn Leu Pro Val Glu Glu  
 625 630 635 640  
 Arg Leu Lys Arg His Ile Ile Asp Gly Glu Arg Leu Gly Leu Glu Glu  
 645 650 655  
 Ala Leu Asn Glu Ala Leu Lys Leu Tyr Ala Pro Leu Asp Ile Ile Asn  
 660 665 670  
 Ile Tyr Leu Leu Asp Gly Met Lys Val Val Gly Glu Leu Phe Gly Ser  
 675 680 685  
 Gly Gln Met Gln Leu Pro Phe Val Leu Gln Ser Ala Gln Thr Met Lys  
 690 695 700

Ala Ala Val Ala Phe Leu Glu Pro His Met Asp Lys Asp Asp Ser Ala  
 705 710 715 720  
 Asp Asn Ala Lys Gly Thr Phe Leu Ile Ala Thr Val Lys Gly Asp Val  
 725 730 735  
 His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly  
 740 745 750  
 Tyr Arg Val Val Asn Leu Gly Ile Lys Gln Pro Val Glu Asn Ile Ile  
 755 760 765  
 Glu Ala Tyr Lys Lys His Arg Pro Asp Cys Ile Ala Met Ser Gly Leu  
 770 775 780  
 Leu Val Lys Ser Thr Ala Phe Met Lys Glu Asn Leu Glu Val Phe Asn  
 785 790 795 800  
 Gln Glu Gly Ile Thr Val Pro Val Ile Leu Gly Gly Ala Ala Leu Thr  
 805 810 815  
 Pro Lys Phe Val His Gln Asp Cys Gln Asn Thr Tyr Lys Gly Gln Val  
 820 825 830  
 Ile Tyr Gly Lys Asp Ala Phe Ala Asp Leu His Phe Met Asp Lys Leu  
 835 840 845  
 Met Pro Ala Lys Asn Ser His Asn Trp Asp Asp Phe Gln Gly Phe Leu  
 850 855 860  
 Gly Glu Tyr Ala Thr Glu Asn Gly His Asn Val Thr Thr Asp Asp Gly  
 865 870 875 880  
 Ala Lys Thr Asn Phe Gly Ile Glu Glu Glu Lys Leu Ile Asp Ala Ser  
 885 890 895  
 Glu Gln Ser Arg Glu Pro Glu Val Ile Asp Thr Val Arg Ser Glu Ala  
 900 905 910  
 Val Asp Pro Asp Leu Glu Arg Pro Val Pro Pro Phe Trp Gly Thr Lys  
 915 920 925  
 Ile Leu Gln Ser Ser Asp Ile Ser Leu Asp Glu Val Phe Pro Leu Leu  
 930 935 940  
 Asp Leu Gln Ala Leu Phe Val Gly Gln Trp Gln Phe Arg Lys Pro Arg  
 945 950 955 960  
 Glu Gln Ser Arg Glu Glu Tyr Glu Gln Phe Leu Ala Glu Lys Val His  
 965 970 975  
 Pro Ile Leu Ala Glu Trp Lys Gly Lys Val Met Ala Glu Asn Leu Leu  
 980 985 990  
 His Pro Thr Val Val Tyr Gly Tyr Phe Pro Cys Gln Ser Gln Gly Asn  
 995 1000 1005  
 Thr Leu Leu Ile Tyr Asp Pro Glu Leu Val Ser Gln Asn Asn Gly Gln  
 1010 1015 1020  
 Ile Pro Pro Asp Ala Thr Ala Ile Ala Lys Phe Glu Phe Pro Arg Gln



27

1025                      1030                      1035                      1040

Lys Ser Gly Arg Arg Leu Cys Ile Ala Asp Phe Phe Ala Ser Lys Glu  
                                  1045                      1050                      1055

Ser Gly Ile Thr Asp Val Phe Pro Leu Gln Ala Val Thr Val Gly Glu  
                                  1060                      1065                      1070

Ile Ala Thr Glu Tyr Ala Arg Lys Leu Phe Ala Gly Asp Asn Tyr Thr  
                                  1075                      1080                      1085

Asp Tyr Leu Tyr Phe His Gly Met Ala Val Gln Met Ala Glu Ala Leu  
                                  1090                      1095                      1100

Ala Glu Trp Thr His Gln Arg Ile Arg Gln Glu Leu Gly Phe Gly His  
                                  1105                      1110                      1115                      1120

Leu Asp Pro Asp Asn Ile Arg Asp Leu Leu Gln Gln Arg Tyr Gln Gly  
                                  1125                      1130                      1135

Ser Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn Met Gln Asp Gln  
                                  1140                      1145                      1150

Tyr Thr Gln Leu Glu Leu Leu Gln Thr Glu Arg Ile Gly Leu Tyr Met  
                                  1155                      1160                      1165

Asp Glu Ser Glu Gln Val Tyr Pro Glu Gln Ser Thr Thr Ala Ile Ile  
                                  1170                      1175                      1180

Ser Tyr His Pro Ala Ala Lys Tyr Phe Ser Ala  
                                  1185                      1190                      1195

<210> 7  
 <211> 3561  
 <212> DNA  
 <213> Prochlorococcus marinus

<220>  
 <221> CDS  
 <222> (1) .. (3558)  
 <223> RCK00830

<400> 7

atg gtt tca ttt aga aat tat tta aat aga gat gat aaa cca att att	48
Met Val Ser Phe Arg Asn Tyr Leu Asn Arg Asp Asp Lys Pro Ile Ile	
1                      5                      10                      15	
att ttc gat ggt ggg aca ggt act tct ttt caa aat tta aat tta tca	96
Ile Phe Asp Gly Gly Thr Gly Thr Ser Phe Gln Asn Leu Asn Leu Ser	
20                      25                      30	
tca cat gat ttt ggt gga gat gat tta gag ggt tgc aat gaa aac tta	144
Ser His Asp Phe Gly Gly Asp Asp Leu Glu Gly Cys Asn Glu Asn Leu	
35                      40                      45	
gtt cta tcc tct cct aat act gtt gaa caa gta cat aat tca ttt ctt	192
Val Leu Ser Ser Pro Asn Thr Val Glu Gln Val His Asn Ser Phe Leu	
50                      55                      60	
gaa gca ggt tgt cat gta att gaa acc aat aca ttt ggt gct tca tct	240
Glu Ala Gly Cys His Val Ile Glu Thr Asn Thr Phe Gly Ala Ser Ser	

28

65	70	75	80	
att gtt tta gac gaa tat agt att tct aat aaa gct tat gaa atc aat				288
Ile Val Leu Asp Glu Tyr Ser Ile Ser Asn Lys Ala Tyr Glu Ile Asn				
85		90	95	
aaa aaa gca gct cag ata gct aaa aaa tgt gca aat tta ttt tca tct				336
Lys Lys Ala Ala Gln Ile Ala Lys Lys Cys Ala Asn Leu Phe Ser Ser				
100		105	110	
att aat act cct aga ttt gtc gct gga tca att ggg cca act aca aaa				384
Ile Asn Thr Pro Arg Phe Val Ala Gly Ser Ile Gly Pro Thr Thr Lys				
115		120	125	
tta cca aca tta ggt cat att agt ttt gat aag ctt aaa gat tca tat				432
Leu Pro Thr Leu Gly His Ile Ser Phe Asp Lys Leu Lys Asp Ser Tyr				
130		135	140	
gaa gaa caa ata aat ggt cta att gac gga ggt att gac ctt cta ttg				480
Glu Glu Gln Ile Asn Gly Leu Ile Asp Gly Gly Ile Asp Leu Leu Leu				
145		150	155	160
att gaa aca tgc caa gat gtt tta caa ata aaa tca gca tta tct gct				528
Ile Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ser Ala Leu Ser Ala				
165		170	175	
tct caa gaa gtt att aaa aac agg aat att gaa tta cca ata atg ata				576
Ser Gln Glu Val Ile Lys Asn Arg Asn Ile Glu Leu Pro Ile Met Ile				
180		185	190	
tcc ata act atg gaa acc aca gga acg atg ctt gtc ggg tca gat ata				624
Ser Ile Thr Met Glu Thr Thr Gly Thr Met Leu Val Gly Ser Asp Ile				
195		200	205	
gct tct gca tta aca ata tta gag cca tac aat att gat att ctg gga				672
Ala Ser Ala Leu Thr Ile Leu Glu Pro Tyr Asn Ile Asp Ile Leu Gly				
210		215	220	
ctg aat tgt gca act ggt cca gtt caa atg aaa gaa cat att aag tat				720
Leu Asn Cys Ala Thr Gly Pro Val Gln Met Lys Glu His Ile Lys Tyr				
225		230	235	240
tta gct gaa aat tca cct ttt gca att agt tgt ata cct aat gca gga				768
Leu Ala Glu Asn Ser Pro Phe Ala Ile Ser Cys Ile Pro Asn Ala Gly				
245		250	255	
tta cct gaa aat ata gga ggt gtt gct cac tat aaa tta act cca ttg				816
Leu Pro Glu Asn Ile Gly Gly Val Ala His Tyr Lys Leu Thr Pro Leu				
260		265	270	
gag ttg aaa atg cag tta atg aac ttt att tat gat ttt aac gta caa				864
Glu Leu Lys Met Gln Leu Met Asn Phe Ile Tyr Asp Phe Asn Val Gln				
275		280	285	
ctt att ggc gga tgt tgt ggt act act cct gaa cat atc aag cat tta				912
Leu Ile Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Lys His Leu				
290		295	300	
tca tca atc att gag gaa ata gtt gat aaa aaa ata aat aaa aga ctt				960
Ser Ser Ile Ile Glu Glu Ile Val Asp Lys Lys Ile Asn Lys Arg Leu				
305		310	315	320

29

cct act gta aaa aca aat ttt gtt cct tca gca gct tct ata tat aac Pro Thr Val Lys Thr Asn Phe Val Pro Ser Ala Ala Ser Ile Tyr Asn 325 330 335	1008
gca gtt cca tat aaa caa gat aac tca ata tta ata gtt gga gaa cgt Ala Val Pro Tyr Lys Gln Asp Asn Ser Ile Leu Ile Val Gly Glu Arg 340 345 350	1056
tta aat gct agt gga tca aaa aaa gta agg gaa tta cta aat gaa gat Leu Asn Ala Ser Gly Ser Lys Lys Val Arg Glu Leu Leu Asn Glu Asp 355 360 365	1104
gat tgg gac ggc ctg cta tca att gct aaa caa cag caa aaa gaa aat Asp Trp Asp Gly Leu Leu Ser Ile Ala Lys Gln Gln Gln Lys Glu Asn 370 375 380	1152
gct cac ata cta gat gtc aat gtt gat tat gta gga aga gat gga gtt Ala His Ile Leu Asp Val Asn Val Asp Tyr Val Gly Arg Asp Gly Val 385 390 395 400	1200
aaa gat atg aaa gaa att acc tca aga tta gtt aca aat ata aat ctt Lys Asp Met Lys Glu Ile Thr Ser Arg Leu Val Thr Asn Ile Asn Leu 405 410 415	1248
cca tta atg ata gat tca aca gaa gca gat aaa atg gaa agt gga tta Pro Leu Met Ile Asp Ser Thr Glu Ala Asp Lys Met Glu Ser Gly Leu 420 425 430	1296
aag act gta gga gga aaa tgc att ata aat tca aca aac tac gaa gat Lys Thr Val Gly Gly Lys Cys Ile Ile Asn Ser Thr Asn Tyr Glu Asp 435 440 445	1344
gga gat gac aga ttt aat cag gtc tta aga ctt gca tta gat tat ggt Gly Asp Asp Arg Phe Asn Gln Val Leu Arg Leu Ala Leu Asp Tyr Gly 450 455 460	1392
gct gga ata gta att gga act att gat gaa gat gga atg gca aga aca Ala Gly Ile Val Ile Gly Thr Ile Asp Glu Asp Gly Met Ala Arg Thr 465 470 475 480	1440
tca cag aaa aaa tat gac att gca aaa aga gca tta att aaa act aga Ser Gln Lys Lys Tyr Asp Ile Ala Lys Arg Ala Leu Ile Lys Thr Arg 485 490 495	1488
tca agt ggc ctc gct gat tat gag ata ttt ttt gat cct cta gca ttg Ser Ser Gly Leu Ala Asp Tyr Glu Ile Phe Phe Asp Pro Leu Ala Leu 500 505 510	1536
cca ata tct act gga att gaa gaa gat aga tta aat gct aaa gca act Pro Ile Ser Thr Gly Ile Glu Glu Asp Arg Leu Asn Ala Lys Ala Thr 515 520 525	1584
att gaa gct ata tca aaa ata aga aaa agc ttt cca gat att cat att Ile Glu Ala Ile Ser Lys Ile Arg Lys Ser Phe Pro Asp Ile His Ile 530 535 540	1632
att tta ggg ata tct aat att agt ttc ggg ctt tca cca tta tca aga Ile Leu Gly Ile Ser Asn Ile Ser Phe Gly Leu Ser Pro Leu Ser Arg 545 550 555 560	1680
att aat cta aat tca ata ttt ctc gat gaa tgt ata aag gca gga tta Ile Asn Leu Asn Ser Ile Phe Leu Asp Glu Cys Ile Lys Ala Gly Leu	1728

30

565	570	575	
gat tca gcg att att gca cca aat aaa ata ttg cct ctt tca aaa ata Asp Ser Ala Ile Ile Ala Pro Asn Lys Ile Leu Pro Leu Ser Lys Ile 580 585 590			1776
tct gcg gaa aca aaa aaa tta tgt tta gat tta att tat gac aga aga Ser Ala Glu Thr Lys Lys Leu Cys Leu Asp Leu Ile Tyr Asp Arg Arg 595 600 605			1824
aat ttc gaa aat gaa ata tgt ata tat gat cca tta gtt gaa cta aca Asn Phe Glu Asn Glu Ile Cys Ile Tyr Asp Pro Leu Val Glu Leu Thr 610 615 620			1872
aaa gca ttc caa gat ata aca atc agt gac ttt aaa aaa gga tct act Lys Ala Phe Gln Asp Ile Thr Ile Ser Asp Phe Lys Lys Gly Ser Thr 625 630 635 640			1920
tca aac aaa aac ctc acc tta gaa gaa aaa ctt aaa aac cat att gta Ser Asn Lys Asn Leu Thr Leu Glu Glu Lys Leu Lys Asn His Ile Val 645 650 655			1968
gat ggg gaa aaa ata ggt tta gaa gaa caa tta aat aat gcg ctt aaa Asp Gly Glu Lys Ile Gly Leu Glu Glu Gln Leu Asn Asn Ala Leu Lys 660 665 670			2016
aag tac aaa cca ctt gaa ata att aat act tat tta tta gat gga atg Lys Tyr Lys Pro Leu Glu Ile Ile Asn Thr Tyr Leu Leu Asp Gly Met 675 680 685			2064
aaa gta gtc ggt gaa cta ttt gga tcc ggc caa atg caa tta cct ttt Lys Val Val Gly Glu Leu Phe Gly Ser Gly Gln Met Gln Leu Pro Phe 690 695 700			2112
gta ttg caa tca gcg gaa aca atg aaa ttt gct gtt tca gtg ctt gaa Val Leu Gln Ser Ala Glu Thr Met Lys Phe Ala Val Ser Val Leu Glu 705 710 715 720			2160
cct cat atg gaa aca gta gat gaa aaa ata tct aac gga aaa tta cta Pro His Met Glu Thr Val Asp Glu Lys Ile Ser Asn Gly Lys Leu Leu 725 730 735			2208
ata gca act gtt aaa gga gat gtt cat gat ata ggt aaa aat tta gtt Ile Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Leu Val 740 745 750			2256
gat ata att ctc tca aat aat ggt ttt gat gta atc aac ctt gga att Asp Ile Ile Leu Ser Asn Asn Gly Phe Asp Val Ile Asn Leu Gly Ile 755 760 765			2304
aag caa gat gtt tca gcg att att gat gca caa aaa aaa cat aaa gca Lys Gln Asp Val Ser Ala Ile Ile Asp Ala Gln Lys Lys His Lys Ala 770 775 780			2352
gac tgt att gct atg agt ggt tta ctt gtt aaa tct aca gca ttt atg Asp Cys Ile Ala Met Ser Gly Leu Leu Val Lys Ser Thr Ala Phe Met 785 790 795 800			2400
aag gat aat tta gaa gca ttt aac aat gct gaa att aat gtt cca gtt Lys Asp Asn Leu Glu Ala Phe Asn Asn Ala Glu Ile Asn Val Pro Val 805 810 815			2448

31

att ctt gga ggt gca gca tta act cca aaa ttt gtg aat gaa gat tgt	2496
Ile Leu Gly Gly Ala Ala Leu Thr Pro Lys Phe Val Asn Glu Asp Cys	
820 825 830	
agt cag ata tat aaa ggt aaa att ttg tat ggg aaa gat gct ttt aca	2544
Ser Gln Ile Tyr Lys Gly Lys Ile Leu Tyr Gly Lys Asp Ala Phe Thr	
835 840 845	
gat tta caa ttt atg aat gac tat atg gat agt aaa aag aag ggc aat	2592
Asp Leu Gln Phe Met Asn Asp Tyr Met Asp Ser Lys Lys Lys Gly Asn	
850 855 860	
tgg tct aat gaa aat ggt ttt act aat act gat gat att caa att aaa	2640
Trp Ser Asn Glu Asn Gly Phe Thr Asn Thr Asp Asp Ile Gln Ile Lys	
865 870 875 880	
tta gct tcc cca agg tct tcc gct aaa gat aaa aat tta aat aaa aat	2688
Leu Ala Ser Pro Arg Ser Ser Ala Lys Asp Lys Asn Leu Asn Lys Asn	
885 890 895	
ttt gaa aaa acc aaa agt att caa tta att gag aat ttt aat aga tct	2736
Phe Glu Lys Thr Lys Ser Ile Gln Leu Ile Glu Asn Phe Asn Arg Ser	
900 905 910	
aat ttt gta gag gaa gag gaa cct ata aag gct cca ttt ttg gga act	2784
Asn Phe Val Glu Glu Glu Glu Pro Ile Lys Ala Pro Phe Leu Gly Thr	
915 920 925	
aga gtt ctt caa gat att gaa ata gac ttt gac aaa cta att ttt tat	2832
Arg Val Leu Gln Asp Ile Glu Ile Asp Phe Asp Lys Leu Ile Phe Tyr	
930 935 940	
cta gat aaa aaa gca tta ttt agt ggt caa tgg caa att aaa aaa aat	2880
Leu Asp Lys Lys Ala Leu Phe Ser Gly Gln Trp Gln Ile Lys Lys Asn	
945 950 955 960	
aaa ggt caa tca gta gaa gaa tac aat aat tat tta gat tca tat gca	2928
Lys Gly Gln Ser Val Glu Glu Tyr Asn Asn Tyr Leu Asp Ser Tyr Ala	
965 970 975	
aat cca tta ctt gaa aaa tgg att aat att att tta gat aaa ggc tta	2976
Asn Pro Leu Leu Glu Lys Trp Ile Asn Ile Ile Leu Asp Lys Gly Leu	
980 985 990	
att tca cca aaa gta gtc tat ggc tac ttc cgt tgc ggg agg aat gat	3024
Ile Ser Pro Lys Val Val Tyr Gly Tyr Phe Arg Cys Gly Arg Asn Asp	
995 1000 1005	
aat agt att tat ctc ttt gat aat gta tca aat aaa aga att tct gaa	3072
Asn Ser Ile Tyr Leu Phe Asp Asn Val Ser Asn Lys Arg Ile Ser Glu	
1010 1015 1020	
ttt aac ttt cct aga caa aaa tcg gga aat aat ctt tgt att gca gat	3120
Phe Asn Phe Pro Arg Gln Lys Ser Gly Asn Asn Leu Cys Ile Ala Asp	
1025 1030 1035 1040	
ttt tac tgt gat ctt aaa aat aat gat cca gta gat ata ttt cca atg	3168
Phe Tyr Cys Asp Leu Lys Asn Asn Asp Pro Val Asp Ile Phe Pro Met	
1045 1050 1055	
caa gca gta aca atg ggg gaa ata gct agc gaa tat tcc caa gaa tta	3216
Gln Ala Val Thr Met Gly Glu Ile Ala Ser Glu Tyr Ser Gln Glu Leu	

32

1060	1065	1070	
ttt aaa gct gat aaa tat agt gat tat tta ata ttt cat ggt tta acc			3264
Phe Lys Ala Asp Lys Tyr Ser Asp Tyr Leu Ile Phe His Gly Leu Thr			
1075	1080	1085	
ggt caa tta gca gaa gct ctt gca gaa tat gtt cat tca ata gta aga			3312
Val Gln Leu Ala Glu Ala Leu Ala Glu Tyr Val His Ser Ile Val Arg			
1090	1095	1100	
att gaa tgc gga ttt aaa tca tat gag cca aac aat aac cgt gat ata			3360
Ile Glu Cys Gly Phe Lys Ser Tyr Glu Pro Asn Asn Asn Arg Asp Ile			
1105	1110	1115	1120
tta gct caa aaa tat aga gga gct aga tac tca ttt ggt tat cca gct			3408
Leu Ala Gln Lys Tyr Arg Gly Ala Arg Tyr Ser Phe Gly Tyr Pro Ala			
1125	1130	1135	
tgt cct aaa gtt tct gat tca aat ata cag tta tca tta ttg gat aca			3456
Cys Pro Lys Val Ser Asp Ser Asn Ile Gln Leu Ser Leu Leu Asp Thr			
1140	1145	1150	
aaa agg att aat tta aca atg gat gaa tca gag caa tta cat cct gaa			3504
Lys Arg Ile Asn Leu Thr Met Asp Glu Ser Glu Gln Leu His Pro Glu			
1155	1160	1165	
caa agt act act gct ata att tca ctt cat tca aaa gca aaa tat ttt			3552
Gln Ser Thr Thr Ala Ile Ile Ser Leu His Ser Lys Ala Lys Tyr Phe			
1170	1175	1180	
agt gcc taa			3561
Ser Ala			
1185			
<210> 8			
<211> 1186			
<212> PRT			
<213> Prochlorococcus marinus			
<400> 8			
Met Val Ser Phe Arg Asn Tyr Leu Asn Arg Asp Asp Lys Pro Ile Ile			
1	5	10	15
Ile Phe Asp Gly Gly Thr Gly Thr Ser Phe Gln Asn Leu Asn Leu Ser			
20	25	30	
Ser His Asp Phe Gly Gly Asp Asp Leu Glu Gly Cys Asn Glu Asn Leu			
35	40	45	
Val Leu Ser Ser Pro Asn Thr Val Glu Gln Val His Asn Ser Phe Leu			
50	55	60	
Glu Ala Gly Cys His Val Ile Glu Thr Asn Thr Phe Gly Ala Ser Ser			
65	70	75	80
Ile Val Leu Asp Glu Tyr Ser Ile Ser Asn Lys Ala Tyr Glu Ile Asn			
85	90	95	
Lys Lys Ala Ala Gln Ile Ala Lys Lys Cys Ala Asn Leu Phe Ser Ser			
100	105	110	

33

Ile Asn Thr Pro Arg Phe Val Ala Gly Ser Ile Gly Pro Thr Thr Lys  
 115 120 125

Leu Pro Thr Leu Gly His Ile Ser Phe Asp Lys Leu Lys Asp Ser Tyr  
 130 135 140

Glu Glu Gln Ile Asn Gly Leu Ile Asp Gly Gly Ile Asp Leu Leu Leu  
 145 150 155 160

Ile Glu Thr Cys Gln Asp Val Leu Gln Ile Lys Ser Ala Leu Ser Ala  
 165 170 175

Ser Gln Glu Val Ile Lys Asn Arg Asn Ile Glu Leu Pro Ile Met Ile  
 180 185 190

Ser Ile Thr Met Glu Thr Thr Gly Thr Met Leu Val Gly Ser Asp Ile  
 195 200 205

Ala Ser Ala Leu Thr Ile Leu Glu Pro Tyr Asn Ile Asp Ile Leu Gly  
 210 215 220

Leu Asn Cys Ala Thr Gly Pro Val Gln Met Lys Glu His Ile Lys Tyr  
 225 230 235 240

Leu Ala Glu Asn Ser Pro Phe Ala Ile Ser Cys Ile Pro Asn Ala Gly  
 245 250 255

Leu Pro Glu Asn Ile Gly Gly Val Ala His Tyr Lys Leu Thr Pro Leu  
 260 265 270

Glu Leu Lys Met Gln Leu Met Asn Phe Ile Tyr Asp Phe Asn Val Gln  
 275 280 285

Leu Ile Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Lys His Leu  
 290 295 300

Ser Ser Ile Ile Glu Glu Ile Val Asp Lys Lys Ile Asn Lys Arg Leu  
 305 310 315 320

Pro Thr Val Lys Thr Asn Phe Val Pro Ser Ala Ala Ser Ile Tyr Asn  
 325 330 335

Ala Val Pro Tyr Lys Gln Asp Asn Ser Ile Leu Ile Val Gly Glu Arg  
 340 345 350

Leu Asn Ala Ser Gly Ser Lys Lys Val Arg Glu Leu Leu Asn Glu Asp  
 355 360 365

Asp Trp Asp Gly Leu Leu Ser Ile Ala Lys Gln Gln Gln Lys Glu Asn  
 370 375 380

Ala His Ile Leu Asp Val Asn Val Asp Tyr Val Gly Arg Asp Gly Val  
 385 390 395 400

Lys Asp Met Lys Glu Ile Thr Ser Arg Leu Val Thr Asn Ile Asn Leu  
 405 410 415

Pro Leu Met Ile Asp Ser Thr Glu Ala Asp Lys Met Glu Ser Gly Leu  
 420 425 430

Lys Thr Val Gly Gly Lys Cys Ile Ile Asn Ser Thr Asn Tyr Glu Asp  
 435 440 445

Gly Asp Asp Arg Phe Asn Gln Val Leu Arg Leu Ala Leu Asp Tyr Gly  
 450 455 460  
 Ala Gly Ile Val Ile Gly Thr Ile Asp Glu Asp Gly Met Ala Arg Thr  
 465 470 475 480  
 Ser Gln Lys Lys Tyr Asp Ile Ala Lys Arg Ala Leu Ile Lys Thr Arg  
 485 490 495  
 Ser Ser Gly Leu Ala Asp Tyr Glu Ile Phe Phe Asp Pro Leu Ala Leu  
 500 505 510  
 Pro Ile Ser Thr Gly Ile Glu Glu Asp Arg Leu Asn Ala Lys Ala Thr  
 515 520 525  
 Ile Glu Ala Ile Ser Lys Ile Arg Lys Ser Phe Pro Asp Ile His Ile  
 530 535 540  
 Ile Leu Gly Ile Ser Asn Ile Ser Phe Gly Leu Ser Pro Leu Ser Arg  
 545 550 555 560  
 Ile Asn Leu Asn Ser Ile Phe Leu Asp Glu Cys Ile Lys Ala Gly Leu  
 565 570 575  
 Asp Ser Ala Ile Ile Ala Pro Asn Lys Ile Leu Pro Leu Ser Lys Ile  
 580 585 590  
 Ser Ala Glu Thr Lys Lys Leu Cys Leu Asp Leu Ile Tyr Asp Arg Arg  
 595 600 605  
 Asn Phe Glu Asn Glu Ile Cys Ile Tyr Asp Pro Leu Val Glu Leu Thr  
 610 615 620  
 Lys Ala Phe Gln Asp Ile Thr Ile Ser Asp Phe Lys Lys Gly Ser Thr  
 625 630 635 640  
 Ser Asn Lys Asn Leu Thr Leu Glu Glu Lys Leu Lys Asn His Ile Val  
 645 650 655  
 Asp Gly Glu Lys Ile Gly Leu Glu Glu Gln Leu Asn Asn Ala Leu Lys  
 660 665 670  
 Lys Tyr Lys Pro Leu Glu Ile Ile Asn Thr Tyr Leu Leu Asp Gly Met  
 675 680 685  
 Lys Val Val Gly Glu Leu Phe Gly Ser Gly Gln Met Gln Leu Pro Phe  
 690 695 700  
 Val Leu Gln Ser Ala Glu Thr Met Lys Phe Ala Val Ser Val Leu Glu  
 705 710 715 720  
 Pro His Met Glu Thr Val Asp Glu Lys Ile Ser Asn Gly Lys Leu Leu  
 725 730 735  
 Ile Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Leu Val  
 740 745 750  
 Asp Ile Ile Leu Ser Asn Asn Gly Phe Asp Val Ile Asn Leu Gly Ile  
 755 760 765  
 Lys Gln Asp Val Ser Ala Ile Ile Asp Ala Gln Lys Lys His Lys Ala



35

770                      775                      780  
 Asp Cys Ile Ala Met Ser Gly Leu Leu Val Lys Ser Thr Ala Phe Met  
 785                      790                      795                      800  
 Lys Asp Asn Leu Glu Ala Phe Asn Asn Ala Glu Ile Asn Val Pro Val  
                     805                      810                      815  
 Ile Leu Gly Gly Ala Ala Leu Thr Pro Lys Phe Val Asn Glu Asp Cys  
                     820                      825                      830  
 Ser Gln Ile Tyr Lys Gly Lys Ile Leu Tyr Gly Lys Asp Ala Phe Thr  
                     835                      840                      845  
 Asp Leu Gln Phe Met Asn Asp Tyr Met Asp Ser Lys Lys Lys Gly Asn  
                     850                      855                      860  
 Trp Ser Asn Glu Asn Gly Phe Thr Asn Thr Asp Asp Ile Gln Ile Lys  
 865                      870                      875                      880  
 Leu Ala Ser Pro Arg Ser Ser Ala Lys Asp Lys Asn Leu Asn Lys Asn  
                     885                      890                      895  
 Phe Glu Lys Thr Lys Ser Ile Gln Leu Ile Glu Asn Phe Asn Arg Ser  
                     900                      905                      910  
 Asn Phe Val Glu Glu Glu Glu Pro Ile Lys Ala Pro Phe Leu Gly Thr  
                     915                      920                      925  
 Arg Val Leu Gln Asp Ile Glu Ile Asp Phe Asp Lys Leu Ile Phe Tyr  
                     930                      935                      940  
 Leu Asp Lys Lys Ala Leu Phe Ser Gly Gln Trp Gln Ile Lys Lys Asn  
 945                      950                      955                      960  
 Lys Gly Gln Ser Val Glu Glu Tyr Asn Asn Tyr Leu Asp Ser Tyr Ala  
                     965                      970                      975  
 Asn Pro Leu Leu Glu Lys Trp Ile Asn Ile Ile Leu Asp Lys Gly Leu  
                     980                      985                      990  
 Ile Ser Pro Lys Val Val Tyr Gly Tyr Phe Arg Cys Gly Arg Asn Asp  
                     995                      1000                      1005  
 Asn Ser Ile Tyr Leu Phe Asp Asn Val Ser Asn Lys Arg Ile Ser Glu  
                     1010                      1015                      1020  
 Phe Asn Phe Pro Arg Gln Lys Ser Gly Asn Asn Leu Cys Ile Ala Asp  
 1025                      1030                      1035                      1040  
 Phe Tyr Cys Asp Leu Lys Asn Asn Asp Pro Val Asp Ile Phe Pro Met  
                     1045                      1050                      1055  
 Gln Ala Val Thr Met Gly Glu Ile Ala Ser Glu Tyr Ser Gln Glu Leu  
                     1060                      1065                      1070  
 Phe Lys Ala Asp Lys Tyr Ser Asp Tyr Leu Ile Phe His Gly Leu Thr  
                     1075                      1080                      1085  
 Val Gln Leu Ala Glu Ala Leu Ala Glu Tyr Val His Ser Ile Val Arg  
                     1090                      1095                      1100

36

Ile Glu Cys Gly Phe Lys Ser Tyr Glu Pro Asn Asn Asn Arg Asp Ile  
 1105 1110 1115 1120

Leu Ala Gln Lys Tyr Arg Gly Ala Arg Tyr Ser Phe Gly Tyr Pro Ala  
 1125 1130 1135

Cys Pro Lys Val Ser Asp Ser Asn Ile Gln Leu Ser Leu Leu Asp Thr  
 1140 1145 1150

Lys Arg Ile Asn Leu Thr Met Asp Glu Ser Glu Gln Leu His Pro Glu  
 1155 1160 1165

Gln Ser Thr Thr Ala Ile Ile Ser Leu His Ser Lys Ala Lys Tyr Phe  
 1170 1175 1180

Ser Ala  
 1185

&lt;210&gt; 9

&lt;211&gt; 3048

&lt;212&gt; DNA

&lt;213&gt; Thermus thermophilus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(3045)

&lt;223&gt; RTT00266

&lt;400&gt; 9

atg cgg gcc tac aag gag gcg gca cgg ggg ctt ctt aag ggc ggg gtg 48  
 Met Arg Ala Tyr Lys Glu Ala Ala Arg Gly Leu Leu Lys Gly Gly Val  
 1 5 10 15

gac ctc atc ctc ttg gag acc gcc cag gac atc ctc cag gtg cgc tgc 96  
 Asp Leu Ile Leu Leu Glu Thr Ala Gln Asp Ile Leu Gln Val Arg Cys  
 20 25 30

gcc gtc ttg gcg gtg cgg gag gcc atg gcc gag gtg ggc cgg gag gtg 144  
 Ala Val Leu Ala Val Arg Glu Ala Met Ala Glu Val Gly Arg Glu Val  
 35 40 45

ccc ctc cag gtc cag gtg acc ttt gag gcc acg ggg acg atg ctc gtg 192  
 Pro Leu Gln Val Gln Val Thr Phe Glu Ala Thr Gly Thr Met Leu Val  
 50 55 60

ggc acg gac gag cag gcg gcc ctg gcc gct ctg gag agc ctc ccc gtg 240  
 Gly Thr Asp Glu Gln Ala Ala Leu Ala Ala Leu Glu Ser Leu Pro Val  
 65 70 75 80

gac gtg gtg ggg atg aac tgc gcc acg ggc ccc gac ctc atg gac agc 288  
 Asp Val Val Gly Met Asn Cys Ala Thr Gly Pro Asp Leu Met Asp Ser  
 85 90 95

aag gtg cgc tac ttc gcc gag cac agc acc cgc ttc gtc tcc tgc ctc 336  
 Lys Val Arg Tyr Phe Ala Glu His Ser Thr Arg Phe Val Ser Cys Leu  
 100 105 110

ccg aac gcg ggc ctg ccc cgg aac gag ggg ggg agg gtg gtc tac gac 384  
 Pro Asn Ala Gly Leu Pro Arg Asn Glu Gly Gly Arg Val Val Tyr Asp  
 115 120 125

37

ctc acc ccc gag gag ctc gcc aag tgg cac ctc aag ttc gtg gcc gag	432
Leu Thr Pro Glu Glu Leu Ala Lys Trp His Leu Lys Phe Val Ala Glu	
130 135 140	
tac ggg gtg aac gcc gtg ggg gga tgc tgc ggc acg ggg ccc gag cac	480
Tyr Gly Val Asn Ala Val Gly Gly Cys Cys Gly Thr Gly Pro Glu His	
145 150 155 160	
ata agg aag gtg gcc gag gcg gtg aag ggg ctc gcc ccg aag cca agg	528
Ile Arg Lys Val Ala Glu Ala Val Lys Gly Leu Ala Pro Lys Pro Arg	
165 170 175	
ccc gaa agc ttc cct ccc cag gtg gcc tcc ttg tac cag gcg gtg tcc	576
Pro Glu Ser Phe Pro Pro Gln Val Ala Ser Leu Tyr Gln Ala Val Ser	
180 185 190	
ctc aag cag gag gcg agc ctt ttc ctc gtg ggg gag agg ctc aac gcc	624
Leu Lys Gln Glu Ala Ser Leu Phe Leu Val Gly Glu Arg Leu Asn Ala	
195 200 205	
acg ggg agc aag cgc ttc cgg gag atg ctc ttc gcg aga gac ctc gag	672
Thr Gly Ser Lys Arg Phe Arg Glu Met Leu Phe Ala Arg Asp Leu Glu	
210 215 220	
ggc atc ctc gcc ctc gcc cgg gag cag gtg gag gag ggg gcc cac gcc	720
Gly Ile Leu Ala Leu Ala Arg Glu Gln Val Glu Glu Gly Ala His Ala	
225 230 235 240	
ctg gac ctc tcc gtg gcc tgg acg ggg cgg gac gag ctt gag gac ctc	768
Leu Asp Leu Ser Val Ala Trp Thr Gly Arg Asp Glu Leu Glu Asp Leu	
245 250 255	
cgg tgg ctc ctt ccc cat ctc gcc acc gcc ctt acc gtc ccc gtc atg	816
Arg Trp Leu Leu Pro His Leu Ala Thr Ala Leu Thr Val Pro Val Met	
260 265 270	
gtg gac tcc acc tcc cct gag gcc atg gag ctc gcc ctc aaa tac ctc	864
Val Asp Ser Thr Ser Pro Glu Ala Met Glu Leu Ala Leu Lys Tyr Leu	
275 280 285	
ccg ggc cgg gtc ctc ctg aac tcc gcc aac ctc gag gat ggc tta gag	912
Pro Gly Arg Val Leu Leu Asn Ser Ala Asn Leu Glu Asp Gly Leu Glu	
290 295 300	
cgc ttt gac cgg gtg gcc tcc ctg gcc aag gcc cac ggg gcg gcc ctc	960
Arg Phe Asp Arg Val Ala Ser Leu Ala Lys Ala His Gly Ala Ala Leu	
305 310 315 320	
gtg gtc ctc gcc att gac gag aag ggg atg gcc aag acc cgg gag gag	1008
Val Val Leu Ala Ile Asp Glu Lys Gly Met Ala Lys Thr Arg Glu Glu	
325 330 335	
aag gtg cgg gtg gcc ctg agg atg tac gag cgc ctc acg gag cac cac	1056
Lys Val Arg Val Ala Leu Arg Met Tyr Glu Arg Leu Thr Glu His His	
340 345 350	
ggc ctc cgc ccc gag gac ctc ctc ttt gac ctc ctt acc ttc ccc atc	1104
Gly Leu Arg Pro Glu Asp Leu Leu Phe Asp Leu Leu Thr Phe Pro Ile	
355 360 365	
acc caa ggg gac gag gag agc cgc cct ctg gcc aag gag acc ctc ctc	1152
Thr Gln Gly Asp Glu Glu Ser Arg Pro Leu Ala Lys Glu Thr Leu Leu	

38

370	375	380	
gcc ata gag gag cta cgg gag agg ctt ccc ggg gtg ggc ttc gtc ctt Ala Ile Glu Glu Leu Arg Glu Arg Leu Pro Gly Val Gly Phe Val Leu 385 390 395 400			1200
cgg gtc tcc aac gtc tcc ttc ggg ctc aag ccc cgg gcg agg cgc gtc Arg Val Ser Asn Val Ser Phe Gly Leu Lys Pro Arg Ala Arg Arg Val 405 410 415			1248
ctg aac tcc gtc ttc ctg gac gag gcg agg aaa cgg ggc ctc acc gcg Leu Asn Ser Val Phe Leu Asp Glu Ala Arg Lys Arg Gly Leu Thr Ala 420 425 430			1296
gcc atc gtg gac gcg ggg aag atc ctc ccc ata agc cag atc ccc gag Ala Ile Val Asp Ala Gly Lys Ile Leu Pro Ile Ser Gln Ile Pro Glu 435 440 445			1344
gag gcc tac gcc ctc gcc tta gac ctc atc tac gac cgc cgc aag gag Glu Ala Tyr Ala Leu Ala Leu Asp Leu Ile Tyr Asp Arg Arg Lys Glu 450 455 460			1392
ggc ttt gac ccc ctc ctc gcc ttc atg gcc tac ttt gag gcc cac aag Gly Phe Asp Pro Leu Leu Ala Phe Met Ala Tyr Phe Glu Ala His Lys 465 470 475 480			1440
gag gac ccg ggg aag agg gag gac gcc ttc ctg gcc ctt ccc ctt ctg Glu Asp Pro Gly Lys Arg Glu Asp Ala Phe Leu Ala Leu Pro Leu Leu 485 490 495			1488
gag agg ctc aag cgc cgc gtg gtg gag ggg agg aag cag ggc ctc gag Glu Arg Leu Lys Arg Arg Val Val Glu Gly Arg Lys Gln Gly Leu Glu 500 505 510			1536
gcc gac ctg gag gag gcc ctg aag gcg ggg cac aag ccc ttg gac ctc Ala Asp Leu Glu Glu Ala Leu Lys Ala Gly His Lys Pro Leu Asp Leu 515 520 525			1584
atc aac ggc ccc ctc ctc gcg ggg atg aag gag gtg ggg gac ctc ttc Ile Asn Gly Pro Leu Leu Ala Gly Met Lys Glu Val Gly Asp Leu Phe 530 535 540			1632
ggg gcg ggg aag atg cag ctc ccc ttc gtc ctc cag gcc gcc gag gtg Gly Ala Gly Lys Met Gln Leu Pro Phe Val Leu Gln Ala Ala Glu Val 545 550 555 560			1680
atg aag cgg gcg gtg gcc tac ctc gag ccc cac atg gag aag aag ggg Met Lys Arg Ala Val Ala Tyr Leu Glu Pro His Met Glu Lys Lys Gly 565 570 575			1728
gag ggc aag ggt acc ctg gtc ctc gcc acc gtc aag ggg gac gtg cac Glu Gly Lys Gly Thr Leu Val Leu Ala Thr Val Lys Gly Asp Val His 580 585 590			1776
gac atc ggc aag aac ctg gtg gac atc atc ctc agc aac aac ggc tac Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly Tyr 595 600 605			1824
cgg gtg gtg aac ctg ggg atc aag gtg ccc att gag gag atc ctg aag Arg Val Val Asn Leu Gly Ile Lys Val Pro Ile Glu Glu Ile Leu Lys 610 615 620			1872

39

gcc gtg gag gcg cac aag ccc cac gcc gtg ggc atg tcg ggc ctc ctg Ala Val Glu Ala His Lys Pro His Ala Val Gly Met Ser Gly Leu Leu 625 630 635 640	1920
gtg aag agc acc ctg gtg atg aag gag aac ctg gag tac atg cgg gat Val Lys Ser Thr Leu Val Met Lys Glu Asn Leu Glu Tyr Met Arg Asp 645 650 655	1968
agg ggc tac acc ctc ccc gtg atc ctg ggc ggg gcc gcc ctc acc cgg Arg Gly Tyr Thr Leu Pro Val Ile Leu Gly Gly Ala Ala Leu Thr Arg 660 665 670	2016
agc tac gtg gag gag ctt aag gcc atc tac ccc aac gtc tac tac gcc Ser Tyr Val Glu Glu Leu Lys Ala Ile Tyr Pro Asn Val Tyr Tyr Ala 675 680 685	2064
gag gac gcc ttt gag ggc tta agg ctc atg gag gag ctc acg ggc cac Glu Asp Ala Phe Glu Gly Leu Arg Leu Met Glu Glu Leu Thr Gly His 690 695 700	2112
gcc cct ccc gag ctc acc cgg aag gcc cca gct agg ccc aag cgg gag Ala Pro Pro Glu Leu Thr Arg Lys Ala Pro Ala Arg Pro Lys Arg Glu 705 710 715 720	2160
gcc ccc aag gtg gcg ccc cgc gct cgg ccc gtg ggg gag gcc ccc gcc Ala Pro Lys Val Ala Pro Arg Ala Arg Pro Val Gly Glu Ala Pro Ala 725 730 735	2208
gtc ccc cgg ccc ccc ttc ttc ggc gtg cgg gtg gag gaa ggc ttg gac Val Pro Arg Pro Pro Phe Phe Gly Val Arg Val Glu Glu Gly Leu Asp 740 745 750	2256
ctc gcc acc atc gcc cac tac gtc aac aag ctc gcc ctc tac cgg ggc Leu Ala Thr Ile Ala His Tyr Val Asn Lys Leu Ala Leu Tyr Arg Gly 755 760 765	2304
cag tgg ggc tac agc cgc aag ggc ttt ccc ggg agg cgt ggc agg ccc Gln Trp Gly Tyr Ser Arg Lys Gly Phe Pro Gly Arg Arg Gly Arg Pro 770 775 780	2352
tgg tgg agc ggg agg cgg agc ctg tct tcc aga ggc tcc tca agg agg Trp Trp Ser Gly Arg Arg Ser Leu Ser Ser Arg Gly Ser Ser Arg Arg 785 790 795 800	2400
cga tgg cgg aag ggt ggc ttg aac cca agg tcc tct acg gct tct tcc Arg Trp Arg Lys Gly Gly Leu Asn Pro Arg Ser Ser Thr Ala Ser Ser 805 810 815	2448
ccg tgg ccc ggg agg gga gga gct tct cgt ctt ctc ccc aga gac ggg Pro Trp Pro Gly Arg Gly Gly Ala Ser Arg Leu Leu Pro Arg Asp Gly 820 825 830	2496
gga ggt gct gga gcg ctt ccg ctt ccc ccg gca aag ggg cgg ggg cct Gly Gly Ala Gly Ala Leu Pro Leu Pro Pro Ala Lys Gly Arg Gly Pro 835 840 845	2544
gag cct cgt gga cta ctt ccg ccc ccg gtt tgc cgc gcc ttt ggg gga Glu Pro Arg Gly Leu Leu Pro Pro Val Cys Arg Ala Phe Gly Gly 850 855 860	2592
cga ggc gga ctg gat gcc caa gga ggc ctt ccg ggc ggg ggc cgg gac Arg Gly Gly Leu Asp Ala Gln Gly Gly Leu Pro Gly Gly Gly Arg Asp 865 870 875 880	2640

40

865	870	875	880	
gtc ctc ggg gtc cag ctc gtc acc atg ggg gag gcc cct tcc cga aag				2688
Val Leu Gly Val Gln Leu Val Thr Met Gly Glu Ala Pro Ser Arg Lys				
	885	890	895	
gcc cag gcc ctc ttt gcg tcc ggg gcc tac cag gac tac ctc ttc gtc				2736
Ala Gln Ala Leu Phe Ala Ser Gly Ala Tyr Gln Asp Tyr Leu Phe Val				
	900	905	910	
cac ggc ttc agc gtg gag atg acc gag gcc ttg gcg gag tac tgg cac				2784
His Gly Phe Ser Val Glu Met Thr Glu Ala Leu Ala Glu Tyr Trp His				
	915	920	925	
aag agg atg cgg cag atg tgg ggc atc gcc cac aag gac gcc acc gag				2832
Lys Arg Met Arg Gln Met Trp Gly Ile Ala His Lys Asp Ala Thr Glu				
	930	935	940	
atc cag aag ctc ttc cag cag ggc tac cag ggg gcc cgc tac tcc ttc				2880
Ile Gln Lys Leu Phe Gln Gln Gly Tyr Gln Gly Ala Arg Tyr Ser Phe				
	945	950	955	960
ggc tac ccc gcc tgc ccg gac ctc gcc gac cag gcc aag ctg gac cgg				2928
Gly Tyr Pro Ala Cys Pro Asp Leu Ala Asp Gln Ala Lys Leu Asp Arg				
	965	970	975	
ctc atg ggc ttc cac cgg gtg ggg gtg cac ctc acg gag aac ttc cag				2976
Leu Met Gly Phe His Arg Val Gly Val His Leu Thr Glu Asn Phe Gln				
	980	985	990	
ctg gag ccg gag cac gcc acc agc gcc ctc gtg gtc cac cac ccc gag				3024
Leu Glu Pro Glu His Ala Thr Ser Ala Leu Val Val His His Pro Glu				
	995	1000	1005	
gcc cgc tac ttc agc gtg gac tag				3048
Ala Arg Tyr Phe Ser Val Asp				
	1010	1015		
<210> 10				
<211> 1015				
<212> PRT				
<213> Thermus thermophilus				
<400> 10				
Met Arg Ala Tyr Lys Glu Ala Ala Arg Gly Leu Leu Lys Gly Gly Val				
1	5	10	15	
Asp Leu Ile Leu Leu Glu Thr Ala Gln Asp Ile Leu Gln Val Arg Cys				
20	25	30		
Ala Val Leu Ala Val Arg Glu Ala Met Ala Glu Val Gly Arg Glu Val				
35	40	45		
Pro Leu Gln Val Gln Val Thr Phe Glu Ala Thr Gly Thr Met Leu Val				
50	55	60		
Gly Thr Asp Glu Gln Ala Ala Leu Ala Ala Leu Glu Ser Leu Pro Val				
65	70	75	80	
Asp Val Val Gly Met Asn Cys Ala Thr Gly Pro Asp Leu Met Asp Ser				
85	90	95		

Lys Val Arg Tyr Phe Ala Glu His Ser Thr Arg Phe Val Ser Cys Leu  
 100 105 110  
 Pro Asn Ala Gly Leu Pro Arg Asn Glu Gly Gly Arg Val Val Tyr Asp  
 115 120 125  
 Leu Thr Pro Glu Glu Leu Ala Lys Trp His Leu Lys Phe Val Ala Glu  
 130 135 140  
 Tyr Gly Val Asn Ala Val Gly Gly Cys Cys Gly Thr Gly Pro Glu His  
 145 150 155 160  
 Ile Arg Lys Val Ala Glu Ala Val Lys Gly Leu Ala Pro Lys Pro Arg  
 165 170 175  
 Pro Glu Ser Phe Pro Pro Gln Val Ala Ser Leu Tyr Gln Ala Val Ser  
 180 185 190  
 Leu Lys Gln Glu Ala Ser Leu Phe Leu Val Gly Glu Arg Leu Asn Ala  
 195 200 205  
 Thr Gly Ser Lys Arg Phe Arg Glu Met Leu Phe Ala Arg Asp Leu Glu  
 210 215 220  
 Gly Ile Leu Ala Leu Ala Arg Glu Gln Val Glu Glu Gly Ala His Ala  
 225 230 235 240  
 Leu Asp Leu Ser Val Ala Trp Thr Gly Arg Asp Glu Leu Glu Asp Leu  
 245 250 255  
 Arg Trp Leu Leu Pro His Leu Ala Thr Ala Leu Thr Val Pro Val Met  
 260 265 270  
 Val Asp Ser Thr Ser Pro Glu Ala Met Glu Leu Ala Leu Lys Tyr Leu  
 275 280 285  
 Pro Gly Arg Val Leu Leu Asn Ser Ala Asn Leu Glu Asp Gly Leu Glu  
 290 295 300  
 Arg Phe Asp Arg Val Ala Ser Leu Ala Lys Ala His Gly Ala Ala Leu  
 305 310 315 320  
 Val Val Leu Ala Ile Asp Glu Lys Gly Met Ala Lys Thr Arg Glu Glu  
 325 330 335  
 Lys Val Arg Val Ala Leu Arg Met Tyr Glu Arg Leu Thr Glu His His  
 340 345 350  
 Gly Leu Arg Pro Glu Asp Leu Leu Phe Asp Leu Leu Thr Phe Pro Ile  
 355 360 365  
 Thr Gln Gly Asp Glu Glu Ser Arg Pro Leu Ala Lys Glu Thr Leu Leu  
 370 375 380  
 Ala Ile Glu Glu Leu Arg Glu Arg Leu Pro Gly Val Gly Phe Val Leu  
 385 390 395 400  
 Arg Val Ser Asn Val Ser Phe Gly Leu Lys Pro Arg Ala Arg Arg Val  
 405 410 415  
 Leu Asn Ser Val Phe Leu Asp Glu Ala Arg Lys Arg Gly Leu Thr Ala

42

420	425	430
Ala Ile Val Asp Ala Gly Lys Ile Leu Pro Ile Ser Gln Ile Pro Glu		
435	440	445
Glu Ala Tyr Ala Leu Ala Leu Asp Leu Ile Tyr Asp Arg Arg Lys Glu		
450	455	460
Gly Phe Asp Pro Leu Leu Ala Phe Met Ala Tyr Phe Glu Ala His Lys		
465	470	480
Glu Asp Pro Gly Lys Arg Glu Asp Ala Phe Leu Ala Leu Pro Leu Leu		
485	490	495
Glu Arg Leu Lys Arg Arg Val Val Glu Gly Arg Lys Gln Gly Leu Glu		
500	505	510
Ala Asp Leu Glu Glu Ala Leu Lys Ala Gly His Lys Pro Leu Asp Leu		
515	520	525
Ile Asn Gly Pro Leu Leu Ala Gly Met Lys Glu Val Gly Asp Leu Phe		
530	535	540
Gly Ala Gly Lys Met Gln Leu Pro Phe Val Leu Gln Ala Ala Glu Val		
545	550	560
Met Lys Arg Ala Val Ala Tyr Leu Glu Pro His Met Glu Lys Lys Gly		
565	570	575
Glu Gly Lys Gly Thr Leu Val Leu Ala Thr Val Lys Gly Asp Val His		
580	585	590
Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser Asn Asn Gly Tyr		
595	600	605
Arg Val Val Asn Leu Gly Ile Lys Val Pro Ile Glu Glu Ile Leu Lys		
610	615	620
Ala Val Glu Ala His Lys Pro His Ala Val Gly Met Ser Gly Leu Leu		
625	630	640
Val Lys Ser Thr Leu Val Met Lys Glu Asn Leu Glu Tyr Met Arg Asp		
645	650	655
Arg Gly Tyr Thr Leu Pro Val Ile Leu Gly Gly Ala Ala Leu Thr Arg		
660	665	670
Ser Tyr Val Glu Glu Leu Lys Ala Ile Tyr Pro Asn Val Tyr Tyr Ala		
675	680	685
Glu Asp Ala Phe Glu Gly Leu Arg Leu Met Glu Glu Leu Thr Gly His		
690	695	700
Ala Pro Pro Glu Leu Thr Arg Lys Ala Pro Ala Arg Pro Lys Arg Glu		
705	710	715
Ala Pro Lys Val Ala Pro Arg Ala Arg Pro Val Gly Glu Ala Pro Ala		
725	730	735
Val Pro Arg Pro Pro Phe Phe Gly Val Arg Val Glu Glu Gly Leu Asp		
740	745	750



43

Leu Ala Thr Ile Ala His Tyr Val Asn Lys Leu Ala Leu Tyr Arg Gly  
 755 760 765  
 Gln Trp Gly Tyr Ser Arg Lys Gly Phe Pro Gly Arg Arg Gly Arg Pro  
 770 775 780  
 Trp Trp Ser Gly Arg Arg Ser Leu Ser Ser Arg Gly Ser Ser Arg Arg  
 785 790 795 800  
 Arg Trp Arg Lys Gly Gly Leu Asn Pro Arg Ser Ser Thr Ala Ser Ser  
 805 810 815  
 Pro Trp Pro Gly Arg Gly Gly Ala Ser Arg Leu Leu Pro Arg Asp Gly  
 820 825 830  
 Gly Gly Ala Gly Ala Leu Pro Leu Pro Pro Ala Lys Gly Arg Gly Pro  
 835 840 845  
 Glu Pro Arg Gly Leu Leu Pro Pro Pro Val Cys Arg Ala Phe Gly Gly  
 850 855 860  
 Arg Gly Gly Leu Asp Ala Gln Gly Gly Leu Pro Gly Gly Gly Arg Asp  
 865 870 875 880  
 Val Leu Gly Val Gln Leu Val Thr Met Gly Glu Ala Pro Ser Arg Lys  
 885 890 895  
 Ala Gln Ala Leu Phe Ala Ser Gly Ala Tyr Gln Asp Tyr Leu Phe Val  
 900 905 910  
 His Gly Phe Ser Val Glu Met Thr Glu Ala Leu Ala Glu Tyr Trp His  
 915 920 925  
 Lys Arg Met Arg Gln Met Trp Gly Ile Ala His Lys Asp Ala Thr Glu  
 930 935 940  
 Ile Gln Lys Leu Phe Gln Gln Gly Tyr Gln Gly Ala Arg Tyr Ser Phe  
 945 950 955 960  
 Gly Tyr Pro Ala Cys Pro Asp Leu Ala Asp Gln Ala Lys Leu Asp Arg  
 965 970 975  
 Leu Met Gly Phe His Arg Val Gly Val His Leu Thr Glu Asn Phe Gln  
 980 985 990  
 Leu Glu Pro Glu His Ala Thr Ser Ala Leu Val Val His His Pro Glu  
 995 1000 1005  
 Ala Arg Tyr Phe Ser Val Asp  
 1010 1015

&lt;210&gt; 11

&lt;211&gt; 3441

&lt;212&gt; DNA

&lt;213&gt; Bacillus halodurans

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (3438)

&lt;223&gt; RHD05550

44

<400> 11  
 atg act aaa tcg ttg ttt gaa caa cag tta gag cga aaa atc gtc atc 48  
 Met Thr Lys Ser Leu Phe Glu Gln Gln Leu Glu Arg Lys Ile Val Ile  
 1 5 10 15

ctt gat ggg gcg atg ggg acc atg tta caa gcc gcg aat cta acc gct 96  
 Leu Asp Gly Ala Met Gly Thr Met Leu Gln Ala Ala Asn Leu Thr Ala  
 20 25 30

gat gac ttt ggc gga gaa gag tat gaa ggg tgt aat gaa tat tta aat 144  
 Asp Asp Phe Gly Gly Glu Glu Tyr Glu Gly Cys Asn Glu Tyr Leu Asn  
 35 40 45

gag acg gcc ccc cat gtc gtt gag gac att cat cgc gca tac tta gag 192  
 Glu Thr Ala Pro His Val Val Glu Asp Ile His Arg Ala Tyr Leu Glu  
 50 55 60

gca gga gca gac gtc att gcg acg aac acg ttc ggg gca aca gat atc 240  
 Ala Gly Ala Asp Val Ile Ala Thr Asn Thr Phe Gly Ala Thr Asp Ile  
 65 70 75 80

gtt ctt gac gat tat gat ctc gga tac aaa gca gag gag tta aac ata 288  
 Val Leu Asp Asp Tyr Asp Leu Gly Tyr Lys Ala Glu Glu Leu Asn Ile  
 85 90 95

tgc gcg gtg aaa atc gct aaa cgt gta gct gaa gag ttt tcc act cca 336  
 Cys Ala Val Lys Ile Ala Lys Arg Val Ala Glu Glu Phe Ser Thr Pro  
 100 105 110

gat tgg cct cga ttc gtt gca ggg gcg atg ggg ccg acg acg aaa tct 384  
 Asp Trp Pro Arg Phe Val Ala Gly Ala Met Gly Pro Thr Thr Lys Ser  
 115 120 125

ctt tcc gtc aca ggg ggc gcg aca ttc gaa caa ctt atc gag tct tat 432  
 Leu Ser Val Thr Gly Gly Ala Thr Phe Glu Gln Leu Ile Glu Ser Tyr  
 130 135 140

cgc cag caa gct aca ggt cta att aaa ggc ggg gcg gat att tta tta 480  
 Arg Gln Gln Ala Thr Gly Leu Ile Lys Gly Gly Ala Asp Ile Leu Leu  
 145 150 155 160

ctc gaa acg agc cag gat atg cga aac gtg aag gcg gct tat tta gga 528  
 Leu Glu Thr Ser Gln Asp Met Arg Asn Val Lys Ala Ala Tyr Leu Gly  
 165 170 175

ctg agc caa gcg caa aaa gag cta gag gtg aaa ctg cct ctc att att 576  
 Leu Ser Gln Ala Gln Lys Glu Leu Glu Val Lys Leu Pro Leu Ile Ile  
 180 185 190

tct gga acg att gaa ccg atg gga aca acg ctc gcc ggc caa aac atc 624  
 Ser Gly Thr Ile Glu Pro Met Gly Thr Thr Leu Ala Gly Gln Asn Ile  
 195 200 205

gag gcg ttc tat ttg tca tta gag cat atg aat ccc gtc gtt gtc ggt 672  
 Glu Ala Phe Tyr Leu Ser Leu Glu His Met Asn Pro Val Val Val Gly  
 210 215 220

ctc aac tgc gct aca gga cca gaa ttt atg cgc gat cac ctc cgt tct 720  
 Leu Asn Cys Ala Thr Gly Pro Glu Phe Met Arg Asp His Leu Arg Ser  
 225 230 235 240

ctt tca gac ctt gcg acc tgc tct gta agc tgt tat ccg aat gct ggg 768

45

Leu Ser Asp Leu Ala Thr Cys Ser Val Ser Cys Tyr Pro Asn Ala Gly	
245 250 255	
tta cct gat gaa gag ggg aac tat cac gaa tcc cca gaa tca tta gca	816
Leu Pro Asp Glu Glu Gly Asn Tyr His Glu Ser Pro Glu Ser Leu Ala	
260 265 270	
gcc aag ctc gca ggt ttt gcg gaa aag ggc tgg ttg aat atg gtt ggt	864
Ala Lys Leu Ala Gly Phe Ala Glu Lys Gly Trp Leu Asn Met Val Gly	
275 280 285	
ggc tgt tgc ggg acg act cca gac cac att cgt gct ctt ttg gac gtt	912
Gly Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Leu Leu Asp Val	
290 295 300	
atg aag caa ttt gag ccg aga caa cca aaa ggg gat cac ccc cac tcg	960
Met Lys Gln Phe Glu Pro Arg Gln Pro Lys Gly Asp His Pro His Ser	
305 310 315 320	
gtc tca gga att gag cca ctg tta tac gat gac agc atg cgt cca cta	1008
Val Ser Gly Ile Glu Pro Leu Leu Tyr Asp Asp Ser Met Arg Pro Leu	
325 330 335	
ttt gtc ggt gaa cgg aca aac gtc atc ggg tct cgt aaa ttt aaa cgg	1056
Phe Val Gly Glu Arg Thr Asn Val Ile Gly Ser Arg Lys Phe Lys Arg	
340 345 350	
ttg atc gaa gaa gaa aaa tat gaa gaa gcc tca gaa att gca aga tcc	1104
Leu Ile Glu Glu Glu Lys Tyr Glu Glu Ala Ser Glu Ile Ala Arg Ser	
355 360 365	
caa gtg aag aaa ggg gcc cac gtt atc gat gtt tgt ctt gct gat ccg	1152
Gln Val Lys Lys Gly Ala His Val Ile Asp Val Cys Leu Ala Asp Pro	
370 375 380	
gat cgc gat gaa atg gag gac atg gag gaa ttt tta aaa ttc gtg atc	1200
Asp Arg Asp Glu Met Glu Asp Met Glu Glu Phe Leu Lys Phe Val Ile	
385 390 395 400	
aac aaa gtg aag gta ccg ctc atg att gac tcc acc gac gaa aag gta	1248
Asn Lys Val Lys Val Pro Leu Met Ile Asp Ser Thr Asp Glu Lys Val	
405 410 415	
att gaa caa gcg ctt acg tat tca caa ggg aaa gcg atc att aat tcg	1296
Ile Glu Gln Ala Leu Thr Tyr Ser Gln Gly Lys Ala Ile Ile Asn Ser	
420 425 430	
atc aac tta gag gac ggc gaa gaa cgt ttt gaa aaa gtg gtc ccg ctc	1344
Ile Asn Leu Glu Asp Gly Glu Glu Arg Phe Glu Lys Val Val Pro Leu	
435 440 445	
gtc cat aag tat gga gcc gcg gtt gtc gtt ggt acg atc gac gaa gaa	1392
Val His Lys Tyr Gly Ala Ala Val Val Val Gly Thr Ile Asp Glu Glu	
450 455 460	
gga atg gcg att acg gca gaa aaa aaa tta gcg gtt gcg aaa cga tca	1440
Gly Met Ala Ile Thr Ala Glu Lys Lys Leu Ala Val Ala Lys Arg Ser	
465 470 475 480	
tac gac ctg ctc gta aac aaa tac aac att cgt ccg agc gat att att	1488
Tyr Asp Leu Leu Val Asn Lys Tyr Asn Ile Arg Pro Ser Asp Ile Ile	
485 490 495	

ttt gat ccg ctc gtg ttc cca gta gga aca ggc gat gag caa tac att	1536
Phe Asp Pro Leu Val Phe Pro Val Gly Thr Gly Asp Glu Gln Tyr Ile	
500 505 510	
ggc tcg gcg aat gag acg gtg gaa gga att agg agg atc aaa gaa gag	1584
Gly Ser Ala Asn Glu Thr Val Glu Gly Ile Arg Arg Ile Lys Glu Glu	
515 520 525	
ctc cct gaa tgt tta acg att ctt gga gtt agt aac gtg tcg ttc ggt	1632
Leu Pro Glu Cys Leu Thr Ile Leu Gly Val Ser Asn Val Ser Phe Gly	
530 535 540	
ctt ccg cct gtc gga aga gag gtg ctg aac gcg gcg tac tta tac cat	1680
Leu Pro Pro Val Gly Arg Glu Val Leu Asn Ala Ala Tyr Leu Tyr His	
545 550 555 560	
tgt aca caa gct ggc ctt gat tac gct atc gtg aac aca gaa aag ctt	1728
Cys Thr Gln Ala Gly Leu Asp Tyr Ala Ile Val Asn Thr Glu Lys Leu	
565 570 575	
gag cgt tat gcc tcg att tct gat gaa gaa aaa gaa ttg tca agg aag	1776
Glu Arg Tyr Ala Ser Ile Ser Asp Glu Glu Lys Glu Leu Ser Arg Lys	
580 585 590	
ctc tta ttt gaa acg aca gat gaa acg ctc gct gag ttc acc gcc ttt	1824
Leu Leu Phe Glu Thr Thr Asp Glu Thr Leu Ala Glu Phe Thr Ala Phe	
595 600 605	
tat cga ggg aaa aaa gca gag aaa aaa gtg gag act tct aat tta act	1872
Tyr Arg Gly Lys Lys Ala Glu Lys Lys Val Glu Thr Ser Asn Leu Thr	
610 615 620	
ttg gaa gag cgg ttg gca aac tac att gtt gaa ggg tca aag gac gga	1920
Leu Glu Glu Arg Leu Ala Asn Tyr Ile Val Glu Gly Ser Lys Asp Gly	
625 630 635 640	
ctg aca gaa gat tta gat aaa gcg ctc gcg aaa tat gat gat ccg ctt	1968
Leu Thr Glu Asp Leu Asp Lys Ala Leu Ala Lys Tyr Asp Asp Pro Leu	
645 650 655	
gat atc att aac ggc ccg ctc atg aat gga atg gac gaa gtc ggt cgt	2016
Asp Ile Ile Asn Gly Pro Leu Met Asn Gly Met Asp Glu Val Gly Arg	
660 665 670	
ttg ttt aac aat aac gag ctt att gtc gct gaa gta ttg caa agc gct	2064
Leu Phe Asn Asn Asn Glu Leu Ile Val Ala Glu Val Leu Gln Ser Ala	
675 680 685	
gag gtt atg aag gct tcc gtc gcc cac ctt gag cca cat atg gaa aag	2112
Glu Val Met Lys Ala Ser Val Ala His Leu Glu Pro His Met Glu Lys	
690 695 700	
aaa gca gac gat cat gga aaa gga aaa atc att ctt gcc acg gtc aag	2160
Lys Ala Asp Asp His Gly Lys Gly Lys Ile Ile Leu Ala Thr Val Lys	
705 710 715 720	
ggc gat gtt cac gat atc ggg aaa aat cta gtg gaa att att ttg agc	2208
Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Glu Ile Ile Leu Ser	
725 730 735	
aat aat ggt ttc cgc atc gtg aac cta gga att aaa gtt acc tct aat	2256

47

Asn Asn Gly Phe Arg Ile Val Asn Leu Gly Ile Lys Val Thr Ser Asn	
740 745 750	
gag ctg att gaa gcg gtg gcg aga gaa aat cca gat gcg att ggc ttg	2304
Glu Leu Ile Glu Ala Val Ala Arg Glu Asn Pro Asp Ala Ile Gly Leu	
755 760 765	
tca ggg ttg ctc gtc aaa tca gca caa caa atg gta ctt acc gcc caa	2352
Ser Gly Leu Leu Val Lys Ser Ala Gln Gln Met Val Leu Thr Ala Gln	
770 775 780	
gat ttg aag caa caa caa att tcc att ccg att tta gtc gga ggc gca	2400
Asp Leu Lys Gln Gln Ile Ser Ile Pro Ile Leu Val Gly Gly Ala	
785 790 795 800	
gcc ctt acg cgg aaa ttt acg aat aca aaa atc gct cca gag tat gat	2448
Ala Leu Thr Arg Lys Phe Thr Asn Thr Lys Ile Ala Pro Glu Tyr Asp	
805 810 815	
ggt ctc gtc gtc tac gcg aag gat gcg atg aac ggg tta gag ctt gcc	2496
Gly Leu Val Val Tyr Ala Lys Asp Ala Met Asn Gly Leu Glu Leu Ala	
820 825 830	
aat aaa tta atg aaa cct gat gaa cga gaa aag cta gcg gtc tcc ctc	2544
Asn Lys Leu Met Lys Pro Asp Glu Arg Glu Lys Leu Ala Val Ser Leu	
835 840 845	
cat gaa gcg aag gag cag gcg aac tcg agg aca caa atg gga gga ggc	2592
His Glu Ala Lys Glu Gln Ala Asn Ser Arg Thr Gln Met Gly Gly Gly	
850 855 860	
gga act gca gtt gcg gta aag ccg act cga tcc cat gtt tcg aca acg	2640
Gly Thr Ala Val Ala Val Lys Pro Thr Arg Ser His Val Ser Thr Thr	
865 870 875 880	
gtg cct gta gcg gtc cca cct gat gtg aag ccg cac att ttg cgc cac	2688
Val Pro Val Ala Val Pro Pro Asp Val Lys Pro His Ile Leu Arg His	
885 890 895	
cat agc att gcc cat tta gag ccg tat att aac atg cag atg ttg tta	2736
His Ser Ile Ala His Leu Glu Pro Tyr Ile Asn Met Gln Met Leu Leu	
900 905 910	
gga cgt cac tta ggc tta caa ggg aaa gtg agc cgc ctg ctt gca gaa	2784
Gly Arg His Leu Gly Leu Gln Gly Lys Val Ser Arg Leu Leu Ala Glu	
915 920 925	
aaa gac gag aag gct ctt gaa tta aaa gaa aaa gtt gat gcg cta ctc	2832
Lys Asp Glu Lys Ala Leu Glu Leu Lys Glu Lys Val Asp Ala Leu Leu	
930 935 940	
acc agg gtg aaa gag gag cag ctc atg gaa gcc cat ggc atg tat cag	2880
Thr Arg Val Lys Glu Glu Gln Leu Met Glu Ala His Gly Met Tyr Gln	
945 950 955 960	
ttt ttt cct gcc cag tcg gat ggg gac gat att gtc att tat gat caa	2928
Phe Phe Pro Ala Gln Ser Asp Gly Asp Asp Ile Val Ile Tyr Asp Gln	
965 970 975	
acg gga aca aat gaa atc gag cga ttc cat ttt ccg cgt cag aat aag	2976
Thr Gly Thr Asn Glu Ile Glu Arg Phe His Phe Pro Arg Gln Asn Lys	
980 985 990	

gag cct tat ctg tgt ctt gcc gat ttc ctt cgc cca gtt tcc agt ggg 3024  
 Glu Pro Tyr Leu Cys Leu Ala Asp Phe Leu Arg Pro Val Ser Ser Gly  
 995 1000 1005

gaa atg gac tat gtt ggc ttc ctt gct gta acc gca gga aaa ggc att 3072  
 Glu Met Asp Tyr Val Gly Phe Leu Ala Val Thr Ala Gly Lys Gly Ile  
 1010 1015 1020

cgt gaa tta ggg gag cag gcg aaa gag gct gga gac tat tta ttc agt 3120  
 Arg Glu Leu Gly Glu Gln Ala Lys Glu Ala Gly Asp Tyr Leu Phe Ser  
 1025 1030 1035 1040

cac tta atc caa gca aca gcc tta gag atg gcg gaa ggg ttt gcc gag 3168  
 His Leu Ile Gln Ala Thr Ala Leu Glu Met Ala Glu Gly Phe Ala Glu  
 1045 1050 1055

cgt gtc cat cag ctc atg cgt gat aag tgg ggg ttt cct gat tcg gct 3216  
 Arg Val His Gln Leu Met Arg Asp Lys Trp Gly Phe Pro Asp Ser Ala  
 1060 1065 1070

gac ttt aca atg gaa gag cgt ttc gct gca aaa tac cgt ggc atc cgt 3264  
 Asp Phe Thr Met Glu Glu Arg Phe Ala Ala Lys Tyr Arg Gly Ile Arg  
 1075 1080 1085

gta tcg ttt ggc tac cct gca tgc cct gac ttg gat gac caa gca aag 3312  
 Val Ser Phe Gly Tyr Pro Ala Cys Pro Asp Leu Asp Asp Gln Ala Lys  
 1090 1095 1100

ttg ttt aag ctg ttg aag cct gga aag atc gga att gag ttg acg gaa 3360  
 Leu Phe Lys Leu Leu Lys Pro Gly Lys Ile Gly Ile Glu Leu Thr Glu  
 1105 1110 1115 1120

ggg ttt atg atg gag cca gaa gcc tcc gtc acc gcg atg gtg ttt gcc 3408  
 Gly Phe Met Met Glu Pro Glu Ala Ser Val Thr Ala Met Val Phe Ala  
 1125 1130 1135

cat cct gag gct cgc tat ttt aat gtt tta tag 3441  
 His Pro Glu Ala Arg Tyr Phe Asn Val Leu  
 1140 1145

<210> 12  
 <211> 1146  
 <212> PRT  
 <213> Bacillus halodurans

<400> 12  
 Met Thr Lys Ser Leu Phe Glu Gln Gln Leu Glu Arg Lys Ile Val Ile  
 1 5 10 15  
 Leu Asp Gly Ala Met Gly Thr Met Leu Gln Ala Ala Asn Leu Thr Ala  
 20 25 30  
 Asp Asp Phe Gly Gly Glu Glu Tyr Glu Gly Cys Asn Glu Tyr Leu Asn  
 35 40 45  
 Glu Thr Ala Pro His Val Val Glu Asp Ile His Arg Ala Tyr Leu Glu  
 50 55 60  
 Ala Gly Ala Asp Val Ile Ala Thr Asn Thr Phe Gly Ala Thr Asp Ile  
 65 70 75 80

Val Leu Asp Asp Tyr Asp Leu Gly Tyr Lys Ala Glu Glu Leu Asn Ile  
                             85                            90                            95  
 Cys Ala Val Lys Ile Ala Lys Arg Val Ala Glu Glu Phe Ser Thr Pro  
                             100                            105                            110  
 Asp Trp Pro Arg Phe Val Ala Gly Ala Met Gly Pro Thr Thr Lys Ser  
                             115                            120                            125  
 Leu Ser Val Thr Gly Gly Ala Thr Phe Glu Gln Leu Ile Glu Ser Tyr  
                             130                            135                            140  
 Arg Gln Gln Ala Thr Gly Leu Ile Lys Gly Gly Ala Asp Ile Leu Leu  
                             145                            150                            155                            160  
 Leu Glu Thr Ser Gln Asp Met Arg Asn Val Lys Ala Ala Tyr Leu Gly  
                             165                            170                            175  
 Leu Ser Gln Ala Gln Lys Glu Leu Glu Val Lys Leu Pro Leu Ile Ile  
                             180                            185                            190  
 Ser Gly Thr Ile Glu Pro Met Gly Thr Thr Leu Ala Gly Gln Asn Ile  
                             195                            200                            205  
 Glu Ala Phe Tyr Leu Ser Leu Glu His Met Asn Pro Val Val Val Gly  
                             210                            215                            220  
 Leu Asn Cys Ala Thr Gly Pro Glu Phe Met Arg Asp His Leu Arg Ser  
                             225                            230                            235                            240  
 Leu Ser Asp Leu Ala Thr Cys Ser Val Ser Cys Tyr Pro Asn Ala Gly  
                             245                            250                            255  
 Leu Pro Asp Glu Glu Gly Asn Tyr His Glu Ser Pro Glu Ser Leu Ala  
                             260                            265                            270  
 Ala Lys Leu Ala Gly Phe Ala Glu Lys Gly Trp Leu Asn Met Val Gly  
                             275                            280                            285  
 Gly Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Leu Leu Asp Val  
                             290                            295                            300  
 Met Lys Gln Phe Glu Pro Arg Gln Pro Lys Gly Asp His Pro His Ser  
                             305                            310                            315                            320  
 Val Ser Gly Ile Glu Pro Leu Leu Tyr Asp Asp Ser Met Arg Pro Leu  
                             325                            330                            335  
 Phe Val Gly Glu Arg Thr Asn Val Ile Gly Ser Arg Lys Phe Lys Arg  
                             340                            345                            350  
 Leu Ile Glu Glu Glu Lys Tyr Glu Glu Ala Ser Glu Ile Ala Arg Ser  
                             355                            360                            365  
 Gln Val Lys Lys Gly Ala His Val Ile Asp Val Cys Leu Ala Asp Pro  
                             370                            375                            380  
 Asp Arg Asp Glu Met Glu Asp Met Glu Glu Phe Leu Lys Phe Val Ile  
                             385                            390                            395                            400  
 Asn Lys Val Lys Val Pro Leu Met Ile Asp Ser Thr Asp Glu Lys Val

50

405	410	415
Ile Glu Gln Ala Leu Thr Tyr Ser Gln Gly Lys Ala Ile Ile Asn Ser		
420	425	430
Ile Asn Leu Glu Asp Gly Glu Glu Arg Phe Glu Lys Val Val Pro Leu		
435	440	445
Val His Lys Tyr Gly Ala Ala Val Val Val Gly Thr Ile Asp Glu Glu		
450	455	460
Gly Met Ala Ile Thr Ala Glu Lys Lys Leu Ala Val Ala Lys Arg Ser		
465	470	475
Tyr Asp Leu Leu Val Asn Lys Tyr Asn Ile Arg Pro Ser Asp Ile Ile		
485	490	495
Phe Asp Pro Leu Val Phe Pro Val Gly Thr Gly Asp Glu Gln Tyr Ile		
500	505	510
Gly Ser Ala Asn Glu Thr Val Glu Gly Ile Arg Arg Ile Lys Glu Glu		
515	520	525
Leu Pro Glu Cys Leu Thr Ile Leu Gly Val Ser Asn Val Ser Phe Gly		
530	535	540
Leu Pro Pro Val Gly Arg Glu Val Leu Asn Ala Ala Tyr Leu Tyr His		
545	550	555
Cys Thr Gln Ala Gly Leu Asp Tyr Ala Ile Val Asn Thr Glu Lys Leu		
565	570	575
Glu Arg Tyr Ala Ser Ile Ser Asp Glu Glu Lys Glu Leu Ser Arg Lys		
580	585	590
Leu Leu Phe Glu Thr Thr Asp Glu Thr Leu Ala Glu Phe Thr Ala Phe		
595	600	605
Tyr Arg Gly Lys Lys Ala Glu Lys Lys Val Glu Thr Ser Asn Leu Thr		
610	615	620
Leu Glu Glu Arg Leu Ala Asn Tyr Ile Val Glu Gly Ser Lys Asp Gly		
625	630	635
Leu Thr Glu Asp Leu Asp Lys Ala Leu Ala Lys Tyr Asp Asp Pro Leu		
645	650	655
Asp Ile Ile Asn Gly Pro Leu Met Asn Gly Met Asp Glu Val Gly Arg		
660	665	670
Leu Phe Asn Asn Asn Glu Leu Ile Val Ala Glu Val Leu Gln Ser Ala		
675	680	685
Glu Val Met Lys Ala Ser Val Ala His Leu Glu Pro His Met Glu Lys		
690	695	700
Lys Ala Asp Asp His Gly Lys Gly Lys Ile Ile Leu Ala Thr Val Lys		
705	710	715
Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Glu Ile Ile Leu Ser		
725	730	735



51

Asn Asn Gly Phe Arg Ile Val Asn Leu Gly Ile Lys Val Thr Ser Asn  
 740 745 750  
 Glu Leu Ile Glu Ala Val Ala Arg Glu Asn Pro Asp Ala Ile Gly Leu  
 755 760 765  
 Ser Gly Leu Leu Val Lys Ser Ala Gln Gln Met Val Leu Thr Ala Gln  
 770 775 780  
 Asp Leu Lys Gln Gln Gln Ile Ser Ile Pro Ile Leu Val Gly Gly Ala  
 785 790 795 800  
 Ala Leu Thr Arg Lys Phe Thr Asn Thr Lys Ile Ala Pro Glu Tyr Asp  
 805 810 815  
 Gly Leu Val Val Tyr Ala Lys Asp Ala Met Asn Gly Leu Glu Leu Ala  
 820 825 830  
 Asn Lys Leu Met Lys Pro Asp Glu Arg Glu Lys Leu Ala Val Ser Leu  
 835 840 845  
 His Glu Ala Lys Glu Gln Ala Asn Ser Arg Thr Gln Met Gly Gly Gly  
 850 855 860  
 Gly Thr Ala Val Ala Val Lys Pro Thr Arg Ser His Val Ser Thr Thr  
 865 870 875 880  
 Val Pro Val Ala Val Pro Pro Asp Val Lys Pro His Ile Leu Arg His  
 885 890 895  
 His Ser Ile Ala His Leu Glu Pro Tyr Ile Asn Met Gln Met Leu Leu  
 900 905 910  
 Gly Arg His Leu Gly Leu Gln Gly Lys Val Ser Arg Leu Leu Ala Glu  
 915 920 925  
 Lys Asp Glu Lys Ala Leu Glu Leu Lys Glu Lys Val Asp Ala Leu Leu  
 930 935 940  
 Thr Arg Val Lys Glu Glu Gln Leu Met Glu Ala His Gly Met Tyr Gln  
 945 950 955 960  
 Phe Phe Pro Ala Gln Ser Asp Gly Asp Asp Ile Val Ile Tyr Asp Gln  
 965 970 975  
 Thr Gly Thr Asn Glu Ile Glu Arg Phe His Phe Pro Arg Gln Asn Lys  
 980 985 990  
 Glu Pro Tyr Leu Cys Leu Ala Asp Phe Leu Arg Pro Val Ser Ser Gly  
 995 1000 1005  
 Glu Met Asp Tyr Val Gly Phe Leu Ala Val Thr Ala Gly Lys Gly Ile  
 1010 1015 1020  
 Arg Glu Leu Gly Glu Gln Ala Lys Glu Ala Gly Asp Tyr Leu Phe Ser  
 1025 1030 1035 1040  
 His Leu Ile Gln Ala Thr Ala Leu Glu Met Ala Glu Gly Phe Ala Glu  
 1045 1050 1055  
 Arg Val His Gln Leu Met Arg Asp Lys Trp Gly Phe Pro Asp Ser Ala  
 1060 1065 1070

Asp Phe Thr Met Glu Glu Arg Phe Ala Ala Lys Tyr Arg Gly Ile Arg  
 1075 1080 1085

Val Ser Phe Gly Tyr Pro Ala Cys Pro Asp Leu Asp Asp Gln Ala Lys  
 1090 1095 1100

Leu Phe Lys Leu Leu Lys Pro Gly Lys Ile Gly Ile Glu Leu Thr Glu  
 1105 1110 1115 1120

Gly Phe Met Met Glu Pro Glu Ala Ser Val Thr Ala Met Val Phe Ala  
 1125 1130 1135

His Pro Glu Ala Arg Tyr Phe Asn Val Leu  
 1140 1145

<210> 13

<211> 3411

<212> DNA

<213> Bacillus stearothermophilus

<220>

<221> CDS

<222> (1) .. (3408)

<223> RBE02044

<400> 13

atg gct aac gtc acc tta gaa cag caa ctg caa aga aaa att ctt gtc 48  
 Met Ala Asn Val Thr Leu Glu Gln Gln Leu Gln Arg Lys Ile Leu Val  
 1 5 10 15

atc gat ggc gcc atg ggc acg atg atc caa agc gcc aac cta tcg gcc 96  
 Ile Asp Gly Ala Met Gly Thr Met Ile Gln Ser Ala Asn Leu Ser Ala  
 20 25 30

gcc gac ttt ggc ggc gag gcg tat gaa ggg tgc aac gaa tat ttg acc 144  
 Ala Asp Phe Gly Gly Glu Ala Tyr Glu Gly Cys Asn Glu Tyr Leu Thr  
 35 40 45

ctc acc gcc ccg cat gtc atc cgc cgc att cat gaa gcg tac cta gaa 192  
 Leu Thr Ala Pro His Val Ile Arg Arg Ile His Glu Ala Tyr Leu Glu  
 50 55 60

gcc ggt gct gat atc att gaa acg aac acg ttc gga gcg aca cgc atc 240  
 Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr Phe Gly Ala Thr Arg Ile  
 65 70 75 80

gtg ctt gac gaa tat ggc ctc ggt cat ttg gcg ctt gag ctg aac atc 288  
 Val Leu Asp Glu Tyr Gly Leu Gly His Leu Ala Leu Glu Leu Asn Ile  
 85 90 95

gaa gcg gcc aaa ctc gcc aaa caa acg gct gag tcg ttc tcc acc ccg 336  
 Glu Ala Ala Lys Leu Ala Lys Gln Thr Ala Glu Ser Phe Ser Thr Pro  
 100 105 110

gac tgg ccg cgc ttt gtc gcc ggt tcg atg ggg ccg acg acg aaa acg 384  
 Asp Trp Pro Arg Phe Val Ala Gly Ser Met Gly Pro Thr Thr Lys Thr  
 115 120 125

ttg tcg gtg aca ggc ggc gca acg ttt gaa gaa ctc gtc gcc gcc tac 432  
 Leu Ser Val Thr Gly Gly Ala Thr Phe Glu Glu Leu Val Ala Ala Tyr

53

130	135	140	
gaa gaa caa gcg cgc gga ctg ctc tta gga ggc gtc gac ctt ctc cta Glu Glu Gln Ala Arg Gly Leu Leu Leu Gly Gly Val Asp Leu Leu Leu 145 150 155 160			480
ctc gag acg tgc caa gat acg ctg aat gtc aaa gcc ggt ttt ctc ggc Leu Glu Thr Cys Gln Asp Thr Leu Asn Val Lys Ala Gly Phe Leu Gly 165 170 175			528
att tcg aag gcg ttt gaa gcg gtc ggc cgc cgc gtg ccg ctc atg att Ile Ser Lys Ala Phe Glu Ala Val Gly Arg Arg Val Pro Leu Met Ile 180 185 190			576
tcc ggc acg atc gaa ccg atg ggc acg acg ctc gcc ggg cag gcg atc Ser Gly Thr Ile Glu Pro Met Gly Thr Thr Leu Ala Gly Gln Ala Ile 195 200 205			624
gat gcg ttt ttc atc tcg gtg cgc cat atg aag ccg atc gcc gtc ggc Asp Ala Phe Phe Ile Ser Val Arg His Met Lys Pro Ile Ala Val Gly 210 215 220			672
tta aac tgc gca acc ggt ccg gag ttt atg acc gac cat ttg cgc acg Leu Asn Cys Ala Thr Gly Pro Glu Phe Met Thr Asp His Leu Arg Thr 225 230 235 240			720
ctc gcc tcg ctc gct gac acg gcg gtc agc tgc tac ccg aac gcc ggt Leu Ala Ser Leu Ala Asp Thr Ala Val Ser Cys Tyr Pro Asn Ala Gly 245 250 255			768
ctg ccg gat gag gaa ggc cac tat cat gaa acg ccg aat atg ctg gca Leu Pro Asp Glu Glu Gly His Tyr His Glu Thr Pro Asn Met Leu Ala 260 265 270			816
gag aaa atc cgc cgc ttt gcc gaa aag gga tgg atc aac atc gtc ggc Glu Lys Ile Arg Arg Phe Ala Glu Lys Gly Trp Ile Asn Ile Val Gly 275 280 285			864
ggg tgt tgc ggc acg acg ccg gat cat atc cgc gcc att gct gaa gcg Gly Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Ile Ala Glu Ala 290 295 300			912
gtg cgt gat ctc ccg ccg cgg gcg att ccg tct tcg ttt gat gtc cac Val Arg Asp Leu Pro Pro Arg Ala Ile Pro Ser Ser Phe Asp Val His 305 310 315 320			960
gcc gtt tcc ggc atc gag gcg ctc atc tat gat gaa acg atg cgc ccg Ala Val Ser Gly Ile Glu Ala Leu Ile Tyr Asp Glu Thr Met Arg Pro 325 330 335			1008
ctc ttt gtc ggc gag cgg aca aac gtg atc ggc tcg cgc aaa ttc aag Leu Phe Val Gly Glu Arg Thr Asn Val Ile Gly Ser Arg Lys Phe Lys 340 345 350			1056
cgc ctc atc gcc gaa ggg aaa tac gaa gaa gcg gcg gaa atc gcc cgc Arg Leu Ile Ala Glu Gly Lys Tyr Glu Glu Ala Ala Glu Ile Ala Arg 355 360 365			1104
gcc caa gtg aaa aac ggc gcc cat gtc atc gac att tgc ctc gcc gac Ala Gln Val Lys Asn Gly Ala His Val Ile Asp Ile Cys Leu Ala Asp 370 375 380			1152

## 54

cca gac cgc gac gaa ctc cat gac atg gag cag ttc gtc cgc gaa gtc Pro Asp Arg Asp Glu Leu His Asp Met Glu Gln Phe Val Arg Glu Val 385 390 395 400	1200
gtg aaa aaa gtg aaa gtg ccg ctt gtc atc gat tcg acc gac gag cgc Val Lys Lys Val Lys Val Pro Leu Val Ile Asp Ser Thr Asp Glu Arg 405 410 415	1248
gtc atc gaa cgc gcc ctt acg tat tcg caa ggg aag gcg atc atc aac Val Ile Glu Arg Ala Leu Thr Tyr Ser Gln Gly Lys Ala Ile Ile Asn 420 425 430	1296
tcg atc aac ctc gaa gat ggc gaa gag cgg ttt gcg aag gtc gtt cct Ser Ile Asn Leu Glu Asp Gly Glu Glu Arg Phe Ala Lys Val Val Pro 435 440 445	1344
ctc ctg cat caa tac ggc gcc gcc gtt gtc gtc ggc acg atc gat gag Leu Leu His Gln Tyr Gly Ala Ala Val Val Val Gly Thr Ile Asp Glu 450 455 460	1392
caa gga atg gcg gtt aca gcc gaa cgg aaa ttg gaa atc gcc ttg cgt Gln Gly Met Ala Val Thr Ala Glu Arg Lys Leu Glu Ile Ala Leu Arg 465 470 475 480	1440
tcg tat gac ttg ctg gtg aac cgc tac ggc gtc ccc gag cgc gac atc Ser Tyr Asp Leu Leu Val Asn Arg Tyr Gly Val Pro Glu Arg Asp Ile 485 490 495	1488
att ttc gac ccg ctc gtc ttc ccg gtc ggc acc ggc gat gag caa tac Ile Phe Asp Pro Leu Val Phe Pro Val Gly Thr Gly Asp Glu Gln Tyr 500 505 510	1536
atc ggc gcg gcg aaa gaa acc att gag ggc atc cgc ctc att aaa gag Ile Gly Ala Ala Lys Glu Thr Ile Glu Gly Ile Arg Leu Ile Lys Glu 515 520 525	1584
cgg ctg cct cat tgc ttg acg atg ctt ggc atc agc aac gtc tcg ttc Arg Leu Pro His Cys Leu Thr Met Leu Gly Ile Ser Asn Val Ser Phe 530 535 540	1632
ggc ttg ccg ccg gcc gga cgc gag gtg ctc aac tcc gtc ttt ttg tac Gly Leu Pro Pro Ala Gly Arg Glu Val Leu Asn Ser Val Phe Leu Tyr 545 550 555 560	1680
cat tgc acg caa gcc ggg ctc gat tac gcc atc gtc aac acc gag aaa His Cys Thr Gln Ala Gly Leu Asp Tyr Ala Ile Val Asn Thr Glu Lys 565 570 575	1728
ttg gag ccg ttc gcc tcg att ccg gaa gag gaa gtg cga atg gct gag Leu Glu Arg Phe Ala Ser Ile Pro Glu Glu Glu Val Arg Met Ala Glu 580 585 590	1776
gca ctt ctt ttt gac aca aac gac gaa aca tta aac gcc ttt atc gaa Ala Leu Leu Phe Asp Thr Asn Asp Glu Thr Leu Asn Ala Phe Ile Glu 595 600 605	1824
ttt tac cga agc aaa atc acc gcc gcc aaa ccg gcg cag acg aac ttg Phe Tyr Arg Ser Lys Ile Thr Ala Ala Lys Pro Ala Gln Thr Asn Leu 610 615 620	1872
agc ttg gaa gag ccg ctc gcc cgc tac gtt att gaa ggg tcg aaa gac Ser Leu Glu Glu Arg Leu Ala Arg Tyr Val Ile Glu Gly Ser Lys Asp 620	1920

		55			
625	630	635	640		
ggg ctc att ctc gat ttg gaa aag gcg ctt gag acc tac tcc gat ccg Gly Leu Ile Leu Asp Leu Glu Lys Ala Leu Glu Thr Tyr Ser Asp Pro 645 650 655				1968	
ctg tcc atc atc aac ggt ccg ctc atg gcc ggc atg gat gaa gtc ggg Leu Ser Ile Ile Asn Gly Pro Leu Met Ala Gly Met Asp Glu Val Gly 660 665 670				2016	
cgg ctg ttc aac aac aac cag ctc atc gtc gct gaa gta ttg caa agc Arg Leu Phe Asn Asn Asn Gln Leu Ile Val Ala Glu Val Leu Gln Ser 675 680 685				2064	
gcg gaa gtg atg aaa gca gcg gtc gcc ttt tta gag ctg tat atg gaa Ala Glu Val Met Lys Ala Ala Val Ala Phe Leu Glu Leu Tyr Met Glu 690 695 700				2112	
aag aaa gaa gga agc aca aaa gga aaa gtc att ctc gcc acc gtc aaa Lys Lys Glu Gly Ser Thr Lys Gly Lys Val Ile Leu Ala Thr Val Lys 705 710 715 720				2160	
ggc gat gtg cat gac atc ggc aaa aac ttg gtc gac atc att tta agc Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser 725 730 735				2208	
aac aac ggc tac gag gtg atc gac ctc ggc att aaa gtc gct ccg cag Asn Asn Gly Tyr Glu Val Ile Asp Leu Gly Ile Lys Val Ala Pro Gln 740 745 750				2256	
caa ctc att gaa gcg gtg cgc gaa cat cag ccg gac atc atc ggg ttg Gln Leu Ile Glu Ala Val Arg Glu His Gln Pro Asp Ile Ile Gly Leu 755 760 765				2304	
tcg ggc ttg ctt gtg aaa tcg gct caa cag atg gtc gtc acc gcc caa Ser Gly Leu Leu Val Lys Ser Ala Gln Gln Met Val Val Thr Ala Gln 770 775 780				2352	
gac ttg cgc caa gcg ggc atc tcg acc ccg att tta gtc ggc ggc gcc Asp Leu Arg Gln Ala Gly Ile Ser Thr Pro Ile Leu Val Gly Gly Ala 785 790 795 800				2400	
gcc ttg acg cgc aaa ttt acg gaa aac aaa atc gcg ccc gag tac gac Ala Leu Thr Arg Lys Phe Thr Glu Asn Lys Ile Ala Pro Glu Tyr Asp 805 810 815				2448	
ggc gtt gtc ttg tac gcg aaa gac gcc atg gac ggg ctc gcc ctt gcc Gly Val Val Leu Tyr Ala Lys Asp Ala Met Asp Gly Leu Ala Leu Ala 820 825 830				2496	
aac caa atc cag cag ggc gag att gac tac aag aaa aaa gaa acg gcc Asn Gln Ile Gln Gln Gly Glu Ile Asp Tyr Lys Lys Lys Glu Thr Ala 835 840 845				2544	
gaa agc gag cca acg cgg caa acg acg gtg gtc aca gcg gtc aaa tcg Glu Ser Glu Pro Thr Arg Gln Thr Thr Val Val Thr Ala Val Lys Ser 850 855 860				2592	
acc gtc tcg acc gac gtt ccc gtc tac atc ccg gcc gat ctc gag cgc Thr Val Ser Thr Asp Val Pro Val Tyr Ile Pro Ala Asp Leu Glu Arg 865 870 875 880				2640	

56

cac gcg ctg cga aat gtg ccg ctt gac cac att ttg ccg tac gtc aac His Ala Leu Arg Asn Val Pro Leu Asp His Ile Leu Pro Tyr Val Asn 885 890 895	2688
tgg caa atg gtg ctc ggc cac cac ctc ggc ttg aaa gga aaa gtg aaa Trp Gln Met Val Leu Gly His His Leu Gly Leu Lys Gly Lys Val Lys 900 905 910	2736
cgg ctg ctt gaa gag aaa gac gaa aaa gcg ttg gcg tta aaa gcg gtc Arg Leu Leu Glu Glu Lys Asp Glu Lys Ala Leu Ala Leu Lys Ala Val 915 920 925	2784
gtc gac gaa ctg ctc gcc gaa gcg aaa gag cgc cgc tgg att cag ccc Val Asp Glu Leu Leu Ala Glu Ala Lys Glu Arg Arg Trp Ile Gln Pro 930 935 940	2832
gcc ggc gtc tac cgc ttc ttc ccg gcg caa agc gac ggc aac cgg gtt Ala Gly Val Tyr Arg Phe Phe Pro Ala Gln Ser Asp Gly Asn Arg Val 945 950 955 960	2880
tac att tac gat ccg act gac ggc aaa aca gtg ctc gag atg ttc gac Tyr Ile Tyr Asp Pro Thr Asp Gly Lys Thr Val Leu Glu Met Phe Asp 965 970 975	2928
ttt ccg cgc caa ccg cgg gcg ccg tat ctt tgc ctc gcc gat tat ttg Phe Pro Arg Gln Pro Arg Ala Pro Tyr Leu Cys Leu Ala Asp Tyr Leu 980 985 990	2976
aaa tcg aaa gaa agc ggc gaa atg gat tac gtc ggt ttg ttc gcc gtc Lys Ser Lys Glu Ser Gly Glu Met Asp Tyr Val Gly Leu Phe Ala Val 995 1000 1005	3024
acc gct ggg cat ggc gtc cgc gaa ctc gcc cag cgc tgg aag gaa gaa Thr Ala Gly His Gly Val Arg Glu Leu Ala Gln Arg Trp Lys Glu Glu 1010 1015 1020	3072
ggc gaa ttt ttg aaa agc cat gcc atc caa gcg ttg gcg ctc gag att Gly Glu Phe Leu Lys Ser His Ala Ile Gln Ala Leu Ala Leu Glu Ile 1025 1030 1035 1040	3120
gcc gaa ggg ttc gcc gaa cga atc cat caa att atg cgc gac cgc tgg Ala Glu Gly Phe Ala Glu Arg Ile His Gln Ile Met Arg Asp Arg Trp 1045 1050 1055	3168
ggc ttc ccg gac gac ccg gat ttc acg atg gaa gag cgc ttc gcc gcc Gly Phe Pro Asp Asp Pro Asp Phe Thr Met Glu Glu Arg Phe Ala Ala 1060 1065 1070	3216
aaa tac cag ggc cag cgc tac tcg ttc ggc tac ccg gcc tgt ccg aac Lys Tyr Gln Gly Gln Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn 1075 1080 1085	3264
ttg gaa gac cag gag aaa ctg ttc cgt ctg ctt cat cca gaa gac atc Leu Glu Asp Gln Glu Lys Leu Phe Arg Leu Leu His Pro Glu Asp Ile 1090 1095 1100	3312
ggc atc cgt ctc acc gac ggc tat atg atg gaa ccc gaa gca tcg gtt Gly Ile Arg Leu Thr Asp Gly Tyr Met Met Glu Pro Glu Ala Ser Val 1105 1110 1115 1120	3360
tcg gcg atc gtc ttc gcc cat ccg gaa gcg cgg tat ttc aat gtg tta Ser Ala Ile Val Phe Ala His Pro Glu Ala Arg Tyr Phe Asn Val Leu	3408

57

1125

1130

1135

taa

3411

&lt;210&gt; 14

&lt;211&gt; 1136

&lt;212&gt; PRT

&lt;213&gt; Bacillus stearothermophilus

&lt;400&gt; 14

Met Ala Asn Val Thr Leu Glu Gln Gln Leu Gln Arg Lys Ile Leu Val  
 1 5 10 15

Ile Asp Gly Ala Met Gly Thr Met Ile Gln Ser Ala Asn Leu Ser Ala  
 20 25 30

Ala Asp Phe Gly Gly Glu Ala Tyr Glu Gly Cys Asn Glu Tyr Leu Thr  
 35 40 45

Leu Thr Ala Pro His Val Ile Arg Arg Ile His Glu Ala Tyr Leu Glu  
 50 55 60

Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr Phe Gly Ala Thr Arg Ile  
 65 70 75 80

Val Leu Asp Glu Tyr Gly Leu Gly His Leu Ala Leu Glu Leu Asn Ile  
 85 90 95

Glu Ala Ala Lys Leu Ala Lys Gln Thr Ala Glu Ser Phe Ser Thr Pro  
 100 105 110

Asp Trp Pro Arg Phe Val Ala Gly Ser Met Gly Pro Thr Thr Lys Thr  
 115 120 125

Leu Ser Val Thr Gly Gly Ala Thr Phe Glu Glu Leu Val Ala Ala Tyr  
 130 135 140

Glu Glu Gln Ala Arg Gly Leu Leu Leu Gly Gly Val Asp Leu Leu Leu  
 145 150 155 160

Leu Glu Thr Cys Gln Asp Thr Leu Asn Val Lys Ala Gly Phe Leu Gly  
 165 170 175

Ile Ser Lys Ala Phe Glu Ala Val Gly Arg Arg Val Pro Leu Met Ile  
 180 185 190

Ser Gly Thr Ile Glu Pro Met Gly Thr Thr Leu Ala Gly Gln Ala Ile  
 195 200 205

Asp Ala Phe Phe Ile Ser Val Arg His Met Lys Pro Ile Ala Val Gly  
 210 215 220

Leu Asn Cys Ala Thr Gly Pro Glu Phe Met Thr Asp His Leu Arg Thr  
 225 230 235 240

Leu Ala Ser Leu Ala Asp Thr Ala Val Ser Cys Tyr Pro Asn Ala Gly  
 245 250 255

Leu Pro Asp Glu Glu Gly His Tyr His Glu Thr Pro Asn Met Leu Ala  
 260 265 270

58

Glu Lys Ile Arg Arg Phe Ala Glu Lys Gly Trp Ile Asn Ile Val Gly  
 275 280 285  
 Gly Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Ile Ala Glu Ala  
 290 295 300  
 Val Arg Asp Leu Pro Pro Arg Ala Ile Pro Ser Ser Phe Asp Val His  
 305 310 315 320  
 Ala Val Ser Gly Ile Glu Ala Leu Ile Tyr Asp Glu Thr Met Arg Pro  
 325 330 335  
 Leu Phe Val Gly Glu Arg Thr Asn Val Ile Gly Ser Arg Lys Phe Lys  
 340 345 350  
 Arg Leu Ile Ala Glu Gly Lys Tyr Glu Glu Ala Ala Glu Ile Ala Arg  
 355 360 365  
 Ala Gln Val Lys Asn Gly Ala His Val Ile Asp Ile Cys Leu Ala Asp  
 370 375 380  
 Pro Asp Arg Asp Glu Leu His Asp Met Glu Gln Phe Val Arg Glu Val  
 385 390 395 400  
 Val Lys Lys Val Lys Val Pro Leu Val Ile Asp Ser Thr Asp Glu Arg  
 405 410 415  
 Val Ile Glu Arg Ala Leu Thr Tyr Ser Gln Gly Lys Ala Ile Ile Asn  
 420 425 430  
 Ser Ile Asn Leu Glu Asp Gly Glu Glu Arg Phe Ala Lys Val Val Pro  
 435 440 445  
 Leu Leu His Gln Tyr Gly Ala Ala Val Val Val Gly Thr Ile Asp Glu  
 450 455 460  
 Gln Gly Met Ala Val Thr Ala Glu Arg Lys Leu Glu Ile Ala Leu Arg  
 465 470 475 480  
 Ser Tyr Asp Leu Leu Val Asn Arg Tyr Gly Val Pro Glu Arg Asp Ile  
 485 490 495  
 Ile Phe Asp Pro Leu Val Phe Pro Val Gly Thr Gly Asp Glu Gln Tyr  
 500 505 510  
 Ile Gly Ala Ala Lys Glu Thr Ile Glu Gly Ile Arg Leu Ile Lys Glu  
 515 520 525  
 Arg Leu Pro His Cys Leu Thr Met Leu Gly Ile Ser Asn Val Ser Phe  
 530 535 540  
 Gly Leu Pro Pro Ala Gly Arg Glu Val Leu Asn Ser Val Phe Leu Tyr  
 545 550 555 560  
 His Cys Thr Gln Ala Gly Leu Asp Tyr Ala Ile Val Asn Thr Glu Lys  
 565 570 575  
 Leu Glu Arg Phe Ala Ser Ile Pro Glu Glu Glu Val Arg Met Ala Glu  
 580 585 590  
 Ala Leu Leu Phe Asp Thr Asn Asp Glu Thr Leu Asn Ala Phe Ile Glu  
 595 600 605



Phe Tyr Arg Ser Lys Ile Thr Ala Ala Lys Pro Ala Gln Thr Asn Leu  
 610 615 620  
 Ser Leu Glu Glu Arg Leu Ala Arg Tyr Val Ile Glu Gly Ser Lys Asp  
 625 630 635 640  
 Gly Leu Ile Leu Asp Leu Glu Lys Ala Leu Glu Thr Tyr Ser Asp Pro  
 645 650 655  
 Leu Ser Ile Ile Asn Gly Pro Leu Met Ala Gly Met Asp Glu Val Gly  
 660 665 670  
 Arg Leu Phe Asn Asn Asn Gln Leu Ile Val Ala Glu Val Leu Gln Ser  
 675 680 685  
 Ala Glu Val Met Lys Ala Ala Val Ala Phe Leu Glu Leu Tyr Met Glu  
 690 695 700  
 Lys Lys Glu Gly Ser Thr Lys Gly Lys Val Ile Leu Ala Thr Val Lys  
 705 710 715 720  
 Gly Asp Val His Asp Ile Gly Lys Asn Leu Val Asp Ile Ile Leu Ser  
 725 730 735  
 Asn Asn Gly Tyr Glu Val Ile Asp Leu Gly Ile Lys Val Ala Pro Gln  
 740 745 750  
 Gln Leu Ile Glu Ala Val Arg Glu His Gln Pro Asp Ile Ile Gly Leu  
 755 760 765  
 Ser Gly Leu Leu Val Lys Ser Ala Gln Gln Met Val Val Thr Ala Gln  
 770 775 780  
 Asp Leu Arg Gln Ala Gly Ile Ser Thr Pro Ile Leu Val Gly Gly Ala  
 785 790 795 800  
 Ala Leu Thr Arg Lys Phe Thr Glu Asn Lys Ile Ala Pro Glu Tyr Asp  
 805 810 815  
 Gly Val Val Leu Tyr Ala Lys Asp Ala Met Asp Gly Leu Ala Leu Ala  
 820 825 830  
 Asn Gln Ile Gln Gln Gly Glu Ile Asp Tyr Lys Lys Lys Glu Thr Ala  
 835 840 845  
 Glu Ser Glu Pro Thr Arg Gln Thr Thr Val Val Thr Ala Val Lys Ser  
 850 855 860  
 Thr Val Ser Thr Asp Val Pro Val Tyr Ile Pro Ala Asp Leu Glu Arg  
 865 870 875 880  
 His Ala Leu Arg Asn Val Pro Leu Asp His Ile Leu Pro Tyr Val Asn  
 885 890 895  
 Trp Gln Met Val Leu Gly His His Leu Gly Leu Lys Gly Lys Val Lys  
 900 905 910  
 Arg Leu Leu Glu Glu Lys Asp Glu Lys Ala Leu Ala Leu Lys Ala Val  
 915 920 925  
 Val Asp Glu Leu Leu Ala Glu Ala Lys Glu Arg Arg Trp Ile Gln Pro

60

930                      935                      940  
 Ala Gly Val Tyr Arg Phe Phe Pro Ala Gln Ser Asp Gly Asn Arg Val  
 945                      950                      955                      960  
 Tyr Ile Tyr Asp Pro Thr Asp Gly Lys Thr Val Leu Glu Met Phe Asp  
                     965                      970                      975  
 Phe Pro Arg Gln Pro Arg Ala Pro Tyr Leu Cys Leu Ala Asp Tyr Leu  
                     980                      985                      990  
 Lys Ser Lys Glu Ser Gly Glu Met Asp Tyr Val Gly Leu Phe Ala Val  
                     995                      1000                      1005  
 Thr Ala Gly His Gly Val Arg Glu Leu Ala Gln Arg Trp Lys Glu Glu  
 1010                      1015                      1020  
 Gly Glu Phe Leu Lys Ser His Ala Ile Gln Ala Leu Ala Leu Glu Ile  
 1025                      1030                      1035                      1040  
 Ala Glu Gly Phe Ala Glu Arg Ile His Gln Ile Met Arg Asp Arg Trp  
                     1045                      1050                      1055  
 Gly Phe Pro Asp Asp Pro Asp Phe Thr Met Glu Glu Arg Phe Ala Ala  
                     1060                      1065                      1070  
 Lys Tyr Gln Gly Gln Arg Tyr Ser Phe Gly Tyr Pro Ala Cys Pro Asn  
                     1075                      1080                      1085  
 Leu Glu Asp Gln Glu Lys Leu Phe Arg Leu Leu His Pro Glu Asp Ile  
 1090                      1095                      1100  
 Gly Ile Arg Leu Thr Asp Gly Tyr Met Met Glu Pro Glu Ala Ser Val  
 1105                      1110                      1115                      1120  
 Ser Ala Ile Val Phe Ala His Pro Glu Ala Arg Tyr Phe Asn Val Leu  
                     1125                      1130                      1135

<210> 15  
 <211> 3681  
 <212> DNA  
 <213> *Vibrio cholerae*

<220>  
 <221> CDS  
 <222> (1)..(3678)  
 <223> RVC04265

<400> 15  
 gtg gga aaa gaa gta aga caa caa ctc gaa cag caa ttg aaa caa cgt 48  
 Val Gly Lys Glu Val Arg Gln Gln Leu Glu Gln Gln Leu Lys Gln Arg  
   1                    5                    10                    15  
 atc cta ctg att gat ggt ggt atg ggt acc atg att cag agt tat aag 96  
 Ile Leu Leu Ile Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Lys  
   20                    25                    30  
 tta caa gag gaa gac tat cgc ggt gca cga ttt gtc gat tgg cac tgt 144

61

Leu Gln Glu Glu Asp Tyr Arg Gly Ala Arg Phe Val Asp Trp His Cys	
35 40 45	
gat ttg aaa gga aat aac gac ctc tta gtg ctt act cag ccg caa att	192
Asp Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Thr Gln Pro Gln Ile	
50 55 60	
att aaa gag att cac tcc gct tac ctt gaa gcg ggg gcg gat att ctt	240
Ile Lys Glu Ile His Ser Ala Tyr Leu Glu Ala Gly Ala Asp Ile Leu	
65 70 75 80	
gag acc aac acc ttt aac tca acc acg att gcc atg gca gac tat gac	288
Glu Thr Asn Thr Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Asp	
85 90 95	
atg caa tcg ctc agt gct gaa att aac ttt gcc gcg gct aag ctt gca	336
Met Gln Ser Leu Ser Ala Glu Ile Asn Phe Ala Ala Ala Lys Leu Ala	
100 105 110	
cgt gaa gtc gcg gat gag tgg acg gct aaa gat cca agt cgg cca cgc	384
Arg Glu Val Ala Asp Glu Trp Thr Ala Lys Asp Pro Ser Arg Pro Arg	
115 120 125	
tat gtg gct ggt gtg ctt ggg cca acc aac cgt act tgc tct att tcg	432
Tyr Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Cys Ser Ile Ser	
130 135 140	
cca gat gtg aac gat cca gga ttt cgt aac gtc act ttt gat ggg ctt	480
Pro Asp Val Asn Asp Pro Gly Phe Arg Asn Val Thr Phe Asp Gly Leu	
145 150 155 160	
gtt gaa gcc tat tcc gaa tcg acg cgc gct ttg atc aaa ggt ggc agc	528
Val Glu Ala Tyr Ser Glu Ser Thr Arg Ala Leu Ile Lys Gly Gly Ser	
165 170 175	
gat ctg atc ctc att gaa acc atc ttc gat aca ctt aac gcc aaa gcc	576
Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala	
180 185 190	
tgt gcg ttt gcg gtc gat agc gta ttt gaa gag ctg ggc atc agc tta	624
Cys Ala Phe Ala Val Asp Ser Val Phe Glu Glu Leu Gly Ile Ser Leu	
195 200 205	
cct gtg atg att tcc ggc acg att acc gat gcc tct ggg cga act ctg	672
Pro Val Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu	
210 215 220	
tca gga cag aca acg gaa gct ttc tac aac gcc ttg cgt cat gta cgg	720
Ser Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ala Leu Arg His Val Arg	
225 230 235 240	
ccg att tcg ttt ggc ttg aac tgt gcg tta ggt cct gat gag ctg cgc	768
Pro Ile Ser Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg	
245 250 255	
cag tac gtg gaa gag ctt tca cgc att tca gaa tgc tat gtt tcc gcg	816
Gln Tyr Val Glu Glu Leu Ser Arg Ile Ser Glu Cys Tyr Val Ser Ala	
260 265 270	
cac cca aat gcc gga ctg ccc aat gcg ttt ggt gaa tac gat ctc tct	864
His Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Ser	
275 280 285	

gcc gag gaa atg gca gaa cat att gcg gaa tgg gca caa gct ggc ttt Ala Glu Glu Met Ala Glu His Ile Ala Glu Trp Ala Gln Ala Gly Phe 290 295 300	912
ttg aat ttg gtc ggt ggt tgc tgt gga act aca cct gag cat atc gcc Leu Asn Leu Val Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala 305 310 315 320	960
gcc att gcc aaa gcc gtc gag ggt gta aaa cca agg gct ctg cca gat Ala Ile Ala Lys Ala Val Glu Gly Val Lys Pro Arg Ala Leu Pro Asp 325 330 335	1008
ctg aaa gta gaa tgt cgt ctc tcg ggt tta gag ccg ctc aat att ggt Leu Lys Val Glu Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly 340 345 350	1056
cct gaa acc ttg ttt gtt aac gtg ggc gaa cgt act aac gtc acc ggt Pro Glu Thr Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly 355 360 365	1104
tct gcg cgt ttt aag cgt tta att aaa gaa gag caa tac gac gaa gcg Ser Ala Arg Phe Lys Arg Leu Ile Lys Glu Glu Gln Tyr Asp Glu Ala 370 375 380	1152
ctc gat gtg gcg cgt gag caa gtc gaa aac ggc gcg cag atc att gat Leu Asp Val Ala Arg Glu Gln Val Glu Asn Gly Ala Gln Ile Ile Asp 385 390 395 400	1200
atc aac atg gat gaa ggc atg ttg gac gcc gag gcg tgt atg gtg cgc Ile Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Cys Met Val Arg 405 410 415	1248
ttt ttg aat cta tgc gcc tct gaa cca gaa ata tcc aaa gtt ccg gtg Phe Leu Asn Leu Cys Ala Ser Glu Pro Glu Ile Ser Lys Val Pro Val 420 425 430	1296
atg gtc gac tcc tct aaa tgg gaa gtc att gaa gcg ggt ctg aaa tgc Met Val Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys Cys 435 440 445	1344
att cag ggt aaa ggc atc gtc aac tct atc tct cta aaa gaa ggg aaa Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Lys 450 455 460	1392
gag aag ttt att gcc caa gcc aaa ttg gtg cgc cgc tac ggt gcc gcg Glu Lys Phe Ile Ala Gln Ala Lys Leu Val Arg Arg Tyr Gly Ala Ala 465 470 475 480	1440
gtg att gtg atg gca ttt gac gaa gtg ggc caa gcc gat acc cgt gag Val Ile Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Arg Glu 485 490 495	1488
cgc aaa tta gag atc tgt cgt cgg gct tac cat att ttg gtc gat gag Arg Lys Leu Glu Ile Cys Arg Arg Ala Tyr His Ile Leu Val Asp Glu 500 505 510	1536
gtg ggc ttc cca ccg gaa gat att att ttt gac ccg aac atc ttt gct Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala 515 520 525	1584
gtt gcg acc gga att gat gag cac aat aac tac gca ctg gat ttc att	1632

63

Val Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile	
530 535 540	
aat gca gtg gcg gac att aag cgt gag ctg ccg cat gcg atg att tct	1680
Asn Ala Val Ala Asp Ile Lys Arg Glu Leu Pro His Ala Met Ile Ser	
545 550 555 560	
ggc ggt gtt tct aac gtt tcc ttc tct ttc cgc ggc aac aac tat gtg	1728
Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn Tyr Val	
565 570 575	
cgt gaa gcg atc cat gct gtt ttc ctt tat cac tgc ttc aaa cac ggc	1776
Arg Glu Ala Ile His Ala Val Phe Leu Tyr His Cys Phe Lys His Gly	
580 585 590	
atg gac atg ggg att gtc aac gca ggg cag ctt gaa atc tac gat aac	1824
Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr Asp Asn	
595 600 605	
gtt ccg ctg aaa ctg cgt gag gca gtg gaa gat gtg atc ctc aat cga	1872
Val Pro Leu Lys Leu Arg Glu Ala Val Glu Asp Val Ile Leu Asn Arg	
610 615 620	
cgt agc gat ggc acg gaa aga ctg ctt gag atc gcc gaa gcg tat cgc	1920
Arg Ser Asp Gly Thr Glu Arg Leu Leu Glu Ile Ala Glu Ala Tyr Arg	
625 630 635 640	
gaa aac agt gtt ggt aaa gaa gag gat gct tct gca tta gag tgg cgc	1968
Glu Asn Ser Val Gly Lys Glu Glu Asp Ala Ser Ala Leu Glu Trp Arg	
645 650 655	
gca tgg cct gtg gct aag cgc cta gag cac gct tta gtc aaa ggc atc	2016
Ala Trp Pro Val Ala Lys Arg Leu Glu His Ala Leu Val Lys Gly Ile	
660 665 670	
acc gaa ttt atc gtc caa gac act gaa gaa gca cgt cag caa gcc agt	2064
Thr Glu Phe Ile Val Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ser	
675 680 685	
aaa cca ctg gaa gtg att gaa ggg ccg ctg atg gat ggt atg aac gtg	2112
Lys Pro Leu Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val	
690 695 700	
gtc ggt gac ttg ttc ggg gaa ggg aaa atg ttc cta ccg caa gtc gta	2160
Val Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val	
705 710 715 720	
aaa tca gcg cgt gtc atg aaa caa gcc gtt gcg tat ctt gag cct ttc	2208
Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe	
725 730 735	
att aat gcg caa aaa agt ggt agc act tca aat ggt aag att ttg ctg	2256
Ile Asn Ala Gln Lys Ser Gly Ser Thr Ser Asn Gly Lys Ile Leu Leu	
740 745 750	
gcg acc gta aaa ggc gat gtg cat gac att ggt aag aac att gtt ggc	2304
Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly	
755 760 765	
gtc gtg ctg cag tgt aat aac ttc gag atc atc gat ctt ggt gtg atg	2352
Val Val Leu Gln Cys Asn Asn Phe Glu Ile Ile Asp Leu Gly Val Met	
770 775 780	

gtg	cct	tgc	gag	cag	atc	ctc	aaa	gtc	gca	cgc	gag	caa	aat	gtc	gat	2400
Val	Pro	Cys	Glu	Gln	Ile	Leu	Lys	Val	Ala	Arg	Glu	Gln	Asn	Val	Asp	
785					790					795					800	
atc	atc	ggg	ctc	tct	ggg	ctt	atc	acg	cgc	tct	ttg	gat	gag	atg	gta	2448
Ile	Ile	Gly	Leu	Ser	Gly	Leu	Ile	Thr	Pro	Ser	Leu	Asp	Glu	Met	Val	
				805					810					815		
cac	gtg	gcg	aaa	gag	atg	gag	cga	caa	ggg	ttt	gaa	ctg	cca	ctt	ttg	2496
His	Val	Ala	Lys	Glu	Met	Glu	Arg	Gln	Gly	Phe	Glu	Leu	Pro	Leu	Leu	
			820					825					830			
att	ggg	ggg	gca	aca	acg	tct	aaa	gcg	cat	act	gcg	gtg	aag	att	gaa	2544
Ile	Gly	Gly	Ala	Thr	Thr	Ser	Lys	Ala	His	Thr	Ala	Val	Lys	Ile	Glu	
		835					840					845				
cag	aat	tat	cat	gcg	cct	gta	gtg	tac	gtg	aat	aac	gcg	tcg	cgc	gcg	2592
Gln	Asn	Tyr	His	Ala	Pro	Val	Val	Tyr	Val	Asn	Asn	Ala	Ser	Arg	Ala	
	850					855					860					
gta	ggg	gtg	tgc	aca	tca	tta	ttg	tct	gat	gaa	cag	cgc	ccc	gga	ttt	2640
Val	Gly	Val	Cys	Thr	Ser	Leu	Leu	Ser	Asp	Glu	Gln	Arg	Pro	Gly	Phe	
865					870					875					880	
atc	gaa	cgt	ttg	gat	ctc	gat	tat	gag	cgc	acg	cgt	gat	cag	cat	gct	2688
Ile	Glu	Arg	Leu	Asp	Leu	Asp	Tyr	Glu	Arg	Thr	Arg	Asp	Gln	His	Ala	
				885					890					895		
cgt	aaa	acg	ccc	aaa	tcg	cgc	cca	gtc	acg	tta	gag	cag	gca	cgt	gct	2736
Arg	Lys	Thr	Pro	Lys	Ser	Arg	Pro	Val	Thr	Leu	Glu	Gln	Ala	Arg	Ala	
			900					905					910			
aat	aaa	gcg	gcg	ctg	gat	tgg	gca	aat	tac	acg	cgc	ccc	gct	cct	gcg	2784
Asn	Lys	Ala	Ala	Leu	Asp	Trp	Ala	Asn	Tyr	Thr	Pro	Pro	Ala	Pro	Ala	
		915					920					925				
aaa	cgc	ggg	gtg	cat	gtg	ttt	gaa	aac	att	gcg	tta	gcc	aca	cta	cgt	2832
Lys	Pro	Gly	Val	His	Val	Phe	Glu	Asn	Ile	Ala	Leu	Ala	Thr	Leu	Arg	
		930				935					940					
cct	tat	atc	gat	tgg	acg	cct	ttt	ttt	atg	act	tgg	tcg	ctt	atg	ggc	2880
Pro	Tyr	Ile	Asp	Trp	Thr	Pro	Phe	Phe	Met	Thr	Trp	Ser	Leu	Met	Gly	
945					950					955					960	
aaa	tac	cct	gcc	att	ttg	gag	cat	gaa	gag	gtc	ggg	gaa	gag	gcc	aaa	2928
Lys	Tyr	Pro	Ala	Ile	Leu	Glu	His	Glu	Glu	Val	Gly	Glu	Glu	Ala	Lys	
				965					970					975		
cgt	ctg	ttt	cat	gat	gcc	aat	gcc	tta	ctt	gat	aaa	gta	gag	cga	gaa	2976
Arg	Leu	Phe	His	Asp	Ala	Asn	Ala	Leu	Leu	Asp	Lys	Val	Glu	Arg	Glu	
			980					985								

65

Ala His Val Leu Tyr Asn Leu Arg Gln Gln Thr Glu Lys Pro Lys Gly  
 1025 1030 1035 1040

gcc aac tac tgt ttg tcg gac tat gtt gct ccg aaa gag agc ggt aaa 3168  
 Ala Asn Tyr Cys Leu Ser Asp Tyr Val Ala Pro Lys Glu Ser Gly Lys  
 1045 1050 1055

cgt gat tgg att ggc gcg ttt gca gta act ggt ggc att ggt gag cga 3216  
 Arg Asp Trp Ile Gly Ala Phe Ala Val Thr Gly Gly Ile Gly Glu Arg  
 1060 1065 1070

gcc ttg gcc gat gct tat aaa gct cag ggt gat gat tac aat gcg atc 3264  
 Ala Leu Ala Asp Ala Tyr Lys Ala Gln Gly Asp Asp Tyr Asn Ala Ile  
 1075 1080 1085

atg atc caa gcg gta gcc gat cgt ttg gcg gaa gcc ttt gcg gaa tat 3312  
 Met Ile Gln Ala Val Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr  
 1090 1095 1100

ctg cat gaa aaa gtg cgt aaa gag att tgg ggt tat gcg agc gat gaa 3360  
 Leu His Glu Lys Val Arg Lys Glu Ile Trp Gly Tyr Ala Ser Asp Glu  
 1105 1110 1115 1120

aat ctc tcc aat gat gac ctg atc cgt gag cgt tat cag ggc att cga 3408  
 Asn Leu Ser Asn Asp Asp Leu Ile Arg Glu Arg Tyr Gln Gly Ile Arg  
 1125 1130 1135

ccc gcg ccg ggg tat ccc gcg tgt cct gag cat acc gag aaa gcg act 3456  
 Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Ala Thr  
 1140 1145 1150

ttg tgg cag atg cta aat gtc gaa gag acc ata ggt atg tca ctg acc 3504  
 Leu Trp Gln Met Leu Asn Val Glu Glu Thr Ile Gly Met Ser Leu Thr  
 1155 1160 1165

aca agc tat gcg atg tgg ccg ggc gct tcg gta tcc ggt tgg tat ttc 3552  
 Thr Ser Tyr Ala Met Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe  
 1170 1175 1180

tcg cat ccc gat tct cgc tat ttt gcg gta gcg cag atc caa cca gat 3600  
 Ser His Pro Asp Ser Arg Tyr Phe Ala Val Ala Gln Ile Gln Pro Asp  
 1185 1190 1195 1200

caa ctg cac agc tac gct gag cgt aaa ggt tgg cgt ttg gaa gaa gct 3648  
 Gln Leu His Ser Tyr Ala Glu Arg Lys Gly Trp Arg Leu Glu Glu Ala  
 1205 1210 1215

gaa aag tgg cta gcg cct aac ctt gat gct taa 3681  
 Glu Lys Trp Leu Ala Pro Asn Leu Asp Ala  
 1220 1225

<210> 16  
 <211> 1226  
 <212> PRT  
 <213> Vibrio cholerae

<400> 16  
 Val Gly Lys Glu Val Arg Gln Gln Leu Glu Gln Gln Leu Lys Gln Arg  
 1 5 10 15

Ile Leu Leu Ile Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Lys

66

20	25	30
Leu Gln Glu Glu Asp Tyr Arg Gly Ala Arg Phe Val Asp Trp His Cys		
35	40	45
Asp Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Thr Gln Pro Gln Ile		
50	55	60
Ile Lys Glu Ile His Ser Ala Tyr Leu Glu Ala Gly Ala Asp Ile Leu		
65	70	75
Glu Thr Asn Thr Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Asp		
85	90	95
Met Gln Ser Leu Ser Ala Glu Ile Asn Phe Ala Ala Ala Lys Leu Ala		
100	105	110
Arg Glu Val Ala Asp Glu Trp Thr Ala Lys Asp Pro Ser Arg Pro Arg		
115	120	125
Tyr Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Cys Ser Ile Ser		
130	135	140
Pro Asp Val Asn Asp Pro Gly Phe Arg Asn Val Thr Phe Asp Gly Leu		
145	150	155
Val Glu Ala Tyr Ser Glu Ser Thr Arg Ala Leu Ile Lys Gly Gly Ser		
165	170	175
Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala		
180	185	190
Cys Ala Phe Ala Val Asp Ser Val Phe Glu Glu Leu Gly Ile Ser Leu		
195	200	205
Pro Val Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu		
210	215	220
Ser Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ala Leu Arg His Val Arg		
225	230	235
Pro Ile Ser Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg		
245	250	255
Gln Tyr Val Glu Glu Leu Ser Arg Ile Ser Glu Cys Tyr Val Ser Ala		
260	265	270
His Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Ser		
275	280	285
Ala Glu Glu Met Ala Glu His Ile Ala Glu Trp Ala Gln Ala Gly Phe		
290	295	300
Leu Asn Leu Val Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala		
305	310	315
Ala Ile Ala Lys Ala Val Glu Gly Val Lys Pro Arg Ala Leu Pro Asp		
325	330	335
Leu Lys Val Glu Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly		
340	345	350



67

Pro Glu Thr Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly  
 355 360 365  
 Ser Ala Arg Phe Lys Arg Leu Ile Lys Glu Glu Gln Tyr Asp Glu Ala  
 370 375 380  
 Leu Asp Val Ala Arg Glu Gln Val Glu Asn Gly Ala Gln Ile Ile Asp  
 385 390 395 400  
 Ile Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Cys Met Val Arg  
 405 410 415  
 Phe Leu Asn Leu Cys Ala Ser Glu Pro Glu Ile Ser Lys Val Pro Val  
 420 425 430  
 Met Val Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys Cys  
 435 440 445  
 Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Lys  
 450 455 460  
 Glu Lys Phe Ile Ala Gln Ala Lys Leu Val Arg Arg Tyr Gly Ala Ala  
 465 470 475 480  
 Val Ile Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Arg Glu  
 485 490 495  
 Arg Lys Leu Glu Ile Cys Arg Arg Ala Tyr His Ile Leu Val Asp Glu  
 500 505 510  
 Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala  
 515 520 525  
 Val Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile  
 530 535 540  
 Asn Ala Val Ala Asp Ile Lys Arg Glu Leu Pro His Ala Met Ile Ser  
 545 550 555 560  
 Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn Tyr Val  
 565 570 575  
 Arg Glu Ala Ile His Ala Val Phe Leu Tyr His Cys Phe Lys His Gly  
 580 585 590  
 Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr Asp Asn  
 595 600 605  
 Val Pro Leu Lys Leu Arg Glu Ala Val Glu Asp Val Ile Leu Asn Arg  
 610 615 620  
 Arg Ser Asp Gly Thr Glu Arg Leu Leu Glu Ile Ala Glu Ala Tyr Arg  
 625 630 635 640  
 Glu Asn Ser Val Gly Lys Glu Glu Asp Ala Ser Ala Leu Glu Trp Arg  
 645 650 655  
 Ala Trp Pro Val Ala Lys Arg Leu Glu His Ala Leu Val Lys Gly Ile  
 660 665 670  
 Thr Glu Phe Ile Val Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ser  
 675 680 685

Lys Pro Leu Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val  
 690 695 700  
 Val Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val  
 705 710 715 720  
 Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe  
 725 730 735  
 Ile Asn Ala Gln Lys Ser Gly Ser Thr Ser Asn Gly Lys Ile Leu Leu  
 740 745 750  
 Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly  
 755 760 765  
 Val Val Leu Gln Cys Asn Asn Phe Glu Ile Ile Asp Leu Gly Val Met  
 770 775 780  
 Val Pro Cys Glu Gln Ile Leu Lys Val Ala Arg Glu Gln Asn Val Asp  
 785 790 795 800  
 Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val  
 805 810 815  
 His Val Ala Lys Glu Met Glu Arg Gln Gly Phe Glu Leu Pro Leu Leu  
 820 825 830  
 Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu  
 835 840 845  
 Gln Asn Tyr His Ala Pro Val Val Tyr Val Asn Asn Ala Ser Arg Ala  
 850 855 860  
 Val Gly Val Cys Thr Ser Leu Leu Ser Asp Glu Gln Arg Pro Gly Phe  
 865 870 875 880  
 Ile Glu Arg Leu Asp Leu Asp Tyr Glu Arg Thr Arg Asp Gln His Ala  
 885 890 895  
 Arg Lys Thr Pro Lys Ser Arg Pro Val Thr Leu Glu Gln Ala Arg Ala  
 900 905 910  
 Asn Lys Ala Ala Leu Asp Trp Ala Asn Tyr Thr Pro Pro Ala Pro Ala  
 915 920 925  
 Lys Pro Gly Val His Val Phe Glu Asn Ile Ala Leu Ala Thr Leu Arg  
 930 935 940  
 Pro Tyr Ile Asp Trp Thr Pro Phe Phe Met Thr Trp Ser Leu Met Gly  
 945 950 955 960  
 Lys Tyr Pro Ala Ile Leu Glu His Glu Glu Val Gly Glu Glu Ala Lys  
 965 970 975  
 Arg Leu Phe His Asp Ala Asn Ala Leu Leu Asp Lys Val Glu Arg Glu  
 980 985 990  
 Gly Leu Leu Lys Ala Ser Gly Met Cys Ala Leu Phe Pro Ala Ala Ser  
 995 1000 1005  
 Val Gly Asp Asp Ile Glu Val Tyr Ser Asp Glu Ser Arg Thr Gln Val

69

1010                      1015                      1020  
 Ala His Val Leu Tyr Asn Leu Arg Gln Gln Thr Glu Lys Pro Lys Gly  
 1025                      1030                      1035                      1040  
 Ala Asn Tyr Cys Leu Ser Asp Tyr Val Ala Pro Lys Glu Ser Gly Lys  
                     1045                      1050                      1055  
 Arg Asp Trp Ile Gly Ala Phe Ala Val Thr Gly Gly Ile Gly Glu Arg  
                     1060                      1065                      1070  
 Ala Leu Ala Asp Ala Tyr Lys Ala Gln Gly Asp Asp Tyr Asn Ala Ile  
                     1075                      1080                      1085  
 Met Ile Gln Ala Val Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr  
                     1090                      1095                      1100  
 Leu His Glu Lys Val Arg Lys Glu Ile Trp Gly Tyr Ala Ser Asp Glu  
 1105                      1110                      1115                      1120  
 Asn Leu Ser Asn Asp Asp Leu Ile Arg Glu Arg Tyr Gln Gly Ile Arg  
                     1125                      1130                      1135  
 Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Ala Thr  
                     1140                      1145                      1150  
 Leu Trp Gln Met Leu Asn Val Glu Glu Thr Ile Gly Met Ser Leu Thr  
                     1155                      1160                      1165  
 Thr Ser Tyr Ala Met Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe  
                     1170                      1175                      1180  
 Ser His Pro Asp Ser Arg Tyr Phe Ala Val Ala Gln Ile Gln Pro Asp  
 1185                      1190                      1195                      1200  
 Gln Leu His Ser Tyr Ala Glu Arg Lys Gly Trp Arg Leu Glu Glu Ala  
                     1205                      1210                      1215  
 Glu Lys Trp Leu Ala Pro Asn Leu Asp Ala  
                     1220                      1225

<210> 17  
 <211> 3822  
 <212> DNA  
 <213> Sinorhizobium meliloti

<220>  
 <221> CDS  
 <222> (1) .. (3819)  
 <223> RSM07338

<400> 17  
 gtg agt aaa tcg ata att ctt tgt cgt ttt cag aac ggg aga tct ccc 48  
 Val Ser Lys Ser Ile Ile Leu Cys Arg Phe Gln Asn Gly Arg Ser Pro  
   1                    5                    10                    15  
 atg tcc gcc gcc gac gcc ctc ttt gga aac gtc tcg ccc aag ccg gat 96  
 Met Ser Ala Ala Asp Ala Leu Phe Gly Asn Val Ser Pro Lys Pro Asp  
                     20                    25                    30  
 ggt tcg gaa gtc ttt cgg cag ctc gcc cag gcg gcg gct gaa cgc atc 144

Gly	Ser	Glu	Val	Phe	Arg	Gln	Leu	Ala	Gln	Ala	Ala	Glu	Arg	Ile		
35				40				45								
ctc	atc	atg	gat	ggc	gcc	atg	gga	acg	gag	atc	cag	cag	ctc	ggg	ttc	192
Leu	Ile	Met	Asp	Gly	Ala	Met	Gly	Thr	Glu	Ile	Gln	Gln	Leu	Gly	Phe	
50				55				60								
gtg	gag	gat	cac	ttc	cgc	ggc	gag	cgc	ttc	ggg	ggc	tgc	gcc	tgc	cat	240
Val	Glu	Asp	His	Phe	Arg	Gly	Glu	Arg	Phe	Gly	Gly	Cys	Ala	Cys	His	
65				70				75				80				
cag	cag	ggc	aac	aac	gac	ctc	ctg	acg	ctc	act	cag	ccg	aag	gcg	atc	288
Gln	Gln	Gly	Asn	Asn	Asp	Leu	Leu	Thr	Leu	Thr	Gln	Pro	Lys	Ala	Ile	
85				90				95								
gag	gat	att	cat	tac	cac	tac	gcc	atc	gcc	ggc	gcc	gat	atc	ctc	gaa	336
Glu	Asp	Ile	His	Tyr	His	Tyr	Ala	Ile	Ala	Gly	Ala	Asp	Ile	Leu	Glu	
100				105				110								
acc	aac	acc	ttc	tcc	tcg	acg	cgg	atc	gcc	cag	gcc	gat	tac	ggc	atg	384
Thr	Asn	Thr	Phe	Ser	Ser	Thr	Arg	Ile	Ala	Gln	Ala	Asp	Tyr	Gly	Met	
115				120				125								
gag	gac	atg	gtc	tac	gat	ctc	aat	cgc	gac	ggc	gcg	cgg	ctg	gcg	cgg	432
Glu	Asp	Met	Val	Tyr	Asp	Leu	Asn	Arg	Asp	Gly	Ala	Arg	Leu	Ala	Arg	
130				135				140								
cga	gcc	gcg	aag	cgg	gcc	gag	gcg	gag	gat	ggc	cgg	cgg	cgc	ttc	gtg	480
Arg	Ala	Ala	Lys	Arg	Ala	Glu	Ala	Glu	Asp	Gly	Arg	Arg	Arg	Phe	Val	
145				150				155				160				
gca	ggc	gcg	ctc	ggc	ccc	acc	aac	cgc	acc	gct	tcg	att	tcg	ccg	gac	528
Ala	Gly	Ala	Leu	Gly	Pro	Thr	Asn	Arg	Thr	Ala	Ser	Ile	Ser	Pro	Asp	
165				170				175								
gtc	aac	aac	ccc	ggc	tat	cga	gcc	gtc	agc	ttc	gac	gat	ctg	agg	ctc	576
Val	Asn	Asn	Pro	Gly	Tyr	Arg	Ala	Val	Ser	Phe	Asp	Asp	Leu	Arg	Leu	
180				185				190								
gcc	tat	gcc	gag	cag	gtg	cgg	ggc	ctc	atc	gac	ggc	ggg	gcc	gac	atc	624
Ala	Tyr	Ala	Glu	Gln	Val	Arg	Gly	Leu	Ile	Asp	Gly	Gly	Ala	Asp	Ile	
195				200				205								
atc	ctg	atc	gag	acg	atc	ttc	gac	acg	ctg	aat	gcc	aag	gcg	gcg	atc	672
Ile	Leu	Ile	Glu	Thr	Ile	Phe	Asp	Thr	Leu	Asn	Ala	Lys	Ala	Ala	Ile	
210				215				220								
ttc	gcg	acg	cag	gaa	gtc	ttt	gcc	gaa	aag	ggc	gtc	cgc	ctt	ccg	gtg	720
Phe	Ala	Thr	Gln	Glu	Val	Phe	Ala	Glu	Lys	Gly	Val	Arg	Leu	Pro	Val	
225				230				235				240				
atg	atc	tcc	gga	acg	atc	acc	gat	ctc	tcc	ggc	cgt	acc	ctc	tcc	ggc	768
Met	Ile	Ser	Gly	Thr	Ile	Thr	Asp	Leu	Ser	Gly	Arg	Thr	Leu	Ser	Gly	
245				250				255								
cag	acg	cct	acg	gcc	ttc	tgg	tat	tcg	gtg	cgc	cat	gcg	gat	ccg	ttt	816
Gln	Thr	Pro	Thr	Ala	Phe	Trp	Tyr	Ser	Val	Arg	His	Ala	Asp	Pro	Phe	
260				265				270								

ata gac gag ctt tgc gcg gtc gcc gac acg ctc gtc tgc gcc tat ccg Ile Asp Glu Leu Ser Ala Val Ala Asp Thr Leu Val Cys Ala Tyr Pro 290 295 300	912
aat gcc ggc ctg ccg aac gag ttc ggc cgc tat gac gaa agc ccc gag Asn Ala Gly Leu Pro Asn Glu Phe Gly Arg Tyr Asp Glu Ser Pro Glu 305 310 315 320	960
cag atg gcg gcg cag gtc gag ggc ttc gcc cgg gac ggt ctc gtc aac Gln Met Ala Ala Gln Val Glu Gly Phe Ala Arg Asp Gly Leu Val Asn 325 330 335	1008
atc gtc ggc ggc tgc tgc ggt tcc acg ccg gcc cat atc cgc gcc att Ile Val Gly Gly Cys Cys Gly Ser Thr Pro Ala His Ile Arg Ala Ile 340 345 350	1056
gcc gaa gcg gtt gcc aaa tat ccg ccg cgc cgg gtg ccc gag atc gat Ala Glu Ala Val Ala Lys Tyr Pro Pro Arg Arg Val Pro Glu Ile Asp 355 360 365	1104
cgc cgc atg cgg ctt tcc ggc ctc gaa ccc ttc acg ctt acc gac gag Arg Arg Met Arg Leu Ser Gly Leu Glu Pro Phe Thr Leu Thr Asp Glu 370 375 380	1152
att ccc ttc gtc aac gtc ggc gaa cgc acc aac gtc acc ggc tgc gcg Ile Pro Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala 385 390 395 400	1200
aag ttc cgc aag ctg atc acc gcc ggg gac tac gcc gcc gca ctc gat Lys Phe Arg Lys Leu Ile Thr Ala Gly Asp Tyr Ala Ala Ala Leu Asp 405 410 415	1248
gtg gcg cgt gat cag gtg gcg aat ggc gcc cag atc atc gac gtc aac Val Ala Arg Asp Gln Val Ala Asn Gly Ala Gln Ile Ile Asp Val Asn 420 425 430	1296
atg gac gaa ggc ctg atc gat tgc aag cag gtg atg gtc gag ttc ctg Met Asp Glu Gly Leu Ile Asp Ser Lys Gln Val Met Val Glu Phe Leu 435 440 445	1344
aac ctc gtc gcc tcc gag ccg gat atc gcc cgt gta ccg gtg atg atc Asn Leu Val Ala Ser Glu Pro Asp Ile Ala Arg Val Pro Val Met Ile 450 455 460	1392
gat tgc tgc aaa tgg gag gtg atc gaa gcc ggg ctc aaa tgc gtc cag Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys Cys Val Gln 465 470 475 480	1440
ggc aag gcg ctg gtg aac tcc atc tgc ctc aag gaa ggc gag gcg gct Gly Lys Ala Leu Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Ala Ala 485 490 495	1488
ttc ctg cac cat gcg cgc ctc gtg cgc gcc tat ggc gcc gcg gtc gtg Phe Leu His His Ala Arg Leu Val Arg Ala Tyr Gly Ala Ala Val Val 500 505 510	1536
gtg atg gcg ttc gac gag aag ggc cag gcc gac acg aaa acc cgc aag Val Met Ala Phe Asp Glu Lys Gly Gln Ala Asp Thr Lys Thr Arg Lys 515 520 525	1584
gtg gaa atc tgc cgg cgg gcc tat ccg ctg ctg acg gaa gag gtt ggc	1632

72

Val Glu Ile Cys Arg Arg Ala Tyr Arg Leu Leu Thr Glu Glu Val Gly	
530 535 540	
ttc ccc ccg gag gac atc atc ttc gac ccg aat atc ttc gcg gtc gcg	1680
Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val Ala	
545 550 555 560	
acc ggc atc gag gag cac aac aat tac ggc gtc gac ttc atc gag gcg	1728
Thr Gly Ile Glu Glu His Asn Asn Tyr Gly Val Asp Phe Ile Glu Ala	
565 570 575	
acg cac gag atc atc gcg gca ctg ccg cat gtc cac gtc tcc ggc ggc	1776
Thr His Glu Ile Ile Ala Ala Leu Pro His Val His Val Ser Gly Gly	
580 585 590	
gtg tcg aac ctc tcc ttt tcc ttc cgc ggc aac gag ccg gtg cgc gag	1824
Val Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg Glu	
595 600 605	
gcg atg cac gcc atc ttc ctt tat cac gcg atc cag gcc ggc atg gac	1872
Ala Met His Ala Ile Phe Leu Tyr His Ala Ile Gln Ala Gly Met Asp	
610 615 620	
atg ggc atc gtc aat gcc gga cag ctc gcc gtc tat gat gcg atc gac	1920
Met Gly Ile Val Asn Ala Gly Gln Leu Ala Val Tyr Asp Ala Ile Asp	
625 630 635 640	
ccg gaa ctg cgc gaa acc tgc gag gac gtg gtg ctc aac cgc ccg gcc	1968
Pro Glu Leu Arg Glu Thr Cys Glu Asp Val Val Leu Asn Arg Arg Ala	
645 650 655	
gat tcg acc gag cgc ctc ctg gag atc gcc gag cgc tat cgc ggg aag	2016
Asp Ser Thr Glu Arg Leu Leu Glu Ile Ala Glu Arg Tyr Arg Gly Lys	
660 665 670	
ggc ggg agc cag ggc aag gag aag gac ctt gcc tgg cgc gaa tgg ccg	2064
Gly Gly Ser Gln Gly Lys Glu Lys Asp Leu Ala Trp Arg Glu Trp Pro	
675 680 685	
gtg gag aag ccg ctc gaa cac gcg ctc gtc aat gga att acc gaa ttt	2112
Val Glu Lys Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Glu Phe	
690 695 700	
atc gaa gcc gat acg gaa gag gcc ccg ctt gcc gcc gag ccg ccg ctg	2160
Ile Glu Ala Asp Thr Glu Glu Ala Arg Leu Ala Ala Glu Arg Pro Leu	
705 710 715 720	
cat gtc atc gaa ggc ccg ctg atg gcc ggg atg aac gtc gtg ggc gat	2208
His Val Ile Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp	
725 730 735	
ctc ttc ggt tcc ggc aag atg ttc ctg ccg cag gtg gtc aag tcc gcc	2256
Leu Phe Gly Ser Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala	
740 745 750	
cgg gtg atg aag cag gcc gtt gcg gtg ctg ctc ccc cat atg gag gag	2304
Arg Val Met Lys Gln Ala Val Ala Val Leu Leu Pro His Met Glu Glu	
755 760 765	
gag aag cgc gcc aat ggc ggc ggc gag gcg cgc gag agt gcc ggc aag	2352
Glu Lys Arg Ala Asn Gly Gly Gly Glu Ala Arg Glu Ser Ala Gly Lys	
770 775 780	

atc ctg atg gcg acc gtc aag ggc gac gtg cac gac atc ggc aag aac Ile Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn 785 790 795 800	2400
atc gtc ggc gtc gtg ctc gcc tgc aac aat tac gag atc atc gac ctc Ile Val Gly Val Val Leu Ala Cys Asn Tyr Glu Ile Ile Asp Leu 805 810 815	2448
ggc gtc atg gtg ccc tcg gct aag atc ctc gaa gtg gcg cgc gaa cag Gly Val Met Val Pro Ser Ala Lys Ile Leu Glu Val Ala Arg Glu Gln 820 825 830	2496
aag gtc gac atc gtc ggt ctt tcc ggc ctc atc acg ccg tcg ctg gac Lys Val Asp Ile Val Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp 835 840 845	2544
gag atg gcg cat gtc gct tcc gag ctc gaa cgg gag ggc ttc gat gtc Glu Met Ala His Val Ala Ser Glu Leu Glu Arg Glu Gly Phe Asp Val 850 855 860	2592
ccg ctg ctg atc ggc ggg gcg acg acc agc cgc gtg cac acg gcc gtg Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val 865 870 875 880	2640
aag atc aat ccg cgt tac agc ctc ggc cag acg gtc tat gtc acc gac Lys Ile Asn Pro Arg Tyr Ser Leu Gly Gln Thr Val Tyr Val Thr Asp 885 890 895	2688
gcc agc cgc gcg gtc ggc gtc gta tcg agc ctg ctc tcg ccg gaa gtc Ala Ser Arg Ala Val Gly Val Val Ser Ser Leu Leu Ser Pro Glu Val 900 905 910	2736
cgc gac tcc tac aag aaa acg gtc cgc gcg gag tat ctg aag gtt gcc Arg Asp Ser Tyr Lys Lys Thr Val Arg Ala Glu Tyr Leu Lys Val Ala 915 920 925	2784
gac gca cat gcc cgc aac gaa gcc gag aag cgc cgt ctg ccg ctt tcc Asp Ala His Ala Arg Asn Glu Ala Glu Lys Arg Arg Leu Pro Leu Ser 930 935 940	2832
cag gcg ccg gcg aat gcc ttt cgg ata gat tgg gac gcc cac cag ccg Gln Ala Arg Ala Asn Ala Phe Arg Ile Asp Trp Asp Ala His Gln Pro 945 950 955 960	2880
aag gtt ccg tcc ttc ctc ggc acg cgt gtt ttc gag gga tgg gac ctc Lys Val Pro Ser Phe Leu Gly Thr Arg Val Phe Glu Gly Trp Asp Leu 965 970 975	2928
gcc gaa ctc gcc cgc tat atc gac tgg acg ccg ttc ttc cag acc tgg Ala Glu Leu Ala Arg Tyr Ile Asp Trp Thr Pro Phe Phe Gln Thr Trp 980 985 990	2976
gag ctg aag ggg gta ttc ccg aaa atc ctc gat gac gaa cgc cag ggg Glu Leu Lys Gly Val Phe Pro Lys Ile Leu Asp Asp Glu Arg Gln Gly 995 1000 1005	3024
gct gcc gct cgc cag ctc ttc gag gat gcg cag gcg atg gtc gaa aag Ala Ala Ala Arg Gln Leu Phe Glu Asp Ala Gln Ala Met Val Glu Lys 1010 1015 1020	3072
atc gtg gcc gag gca tgg ttc gcc ccg aag gcc gtg atc ggc ttc tgg	3120

Ile Val Ala Glu Ala	Trp Phe Ala Pro Lys	Ala Val Ile Gly Phe Trp	
1025	1030	1035	1040
cgc gcc gcc agc atg ggc gac gac gtc cgc ctg ttt gcc gac gag gtg			3168
Pro Ala Ala Ser Met	Gly Asp Asp Val Arg	Leu Phe Ala Asp Glu Val	
1045	1050	1055	
cgc gaa gcc gag ctt gcc acc ttc ttc acg ctc cgc cag cag atg gtg			3216
Arg Glu Ala Glu Leu Ala Thr Phe Phe Thr	Leu Arg Gln Gln Met Val		
1060	1065	1070	
aag cgc gac ggc cgg ccg aac gtc gcc ctt gcc gac ttc gtc gcc ccg			3264
Lys Arg Asp Gly Arg Pro Asn Val Ala Leu Ala Asp Phe Val Ala Pro			
1075	1080	1085	
gcg gcg agc ggc aag cgg gac tat gtc ggc ggt ttc gtg gtg acg gcc			3312
Ala Ala Ser Gly Lys Arg Asp Tyr Val Gly Gly Phe Val Val Thr Ala			
1090	1095	1100	
ggc atc gag gaa gtg gcg atc gcc gaa cgc ttc gaa cgg gcg aac gac			3360
Gly Ile Glu Glu Val Ala Ile Ala Glu Arg Phe Glu Arg Ala Asn Asp			
1105	1110	1115	1120
gat tat tcc tcg atc atg gtc aag gcg ctt gcg gac cgc ttc gca gag			3408
Asp Tyr Ser Ser Ile Met Val Lys Ala Leu Ala Asp Arg Phe Ala Glu			
1125	1130	1135	
gcc ttt gcc gag cgc atg cat gaa tat gtc cgc aag gag ctc tgg ggc			3456
Ala Phe Ala Glu Arg Met His Glu Tyr Val Arg Lys Glu Leu Trp Gly			
1140	1145	1150	
tat gct ccg gac gaa gcc ttc acg ccg cag gaa ttg atc gcc gag ccc			3504
Tyr Ala Pro Asp Glu Ala Phe Thr Pro Gln Glu Leu Ile Ala Glu Pro			
1155	1160	1165	
tat gcc ggc atc cgc cct gcg ccc ggc tac ccg gcg cag ccc gac cac			3552
Tyr Ala Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His			
1170	1175	1180	
acg gaa aag gag acg ctt ttc cgg ctc ctg gat gcg gaa gcc gct atc			3600
Thr Glu Lys Glu Thr Leu Phe Arg Leu Leu Asp Ala Glu Ala Ala Ile			
1185	1190	1195	1200
ggc gtc cgg ctc acc gag agc tat gcg atg tgg ccg ggc tct tcg gta			3648
Gly Val Arg Leu Thr Glu Ser Tyr Ala Met Trp Pro Gly Ser Ser Val			
1205	1210	1215	
tcg ggc ctc tat gtc ggc cac ccc gat tcc tat tac ttc ggc gtc gca			3696
Ser Gly Leu Tyr Val Gly His Pro Asp Ser Tyr Tyr Phe Gly Val Ala			
1220	1225	1230	
aag atc gag cgc gat cag gtg gag gac tat gcc gat cgc aag cgc atg			3744
Lys Ile Glu Arg Asp Gln Val Glu Asp Tyr Ala Asp Arg Lys Arg Met			
1235	1240	1245	
agc gtc cgc gag gtc gag cgc tgg ctt tcg ccg atc ctc aat tac gtg			3792
Ser Val Arg Glu Val Glu Arg Trp Leu Ser Pro Ile Leu Asn Tyr Val			
1250	1255	1260	
ccg atg ccg gag acg gaa gcg gcg gag tag			3822
Pro Met Pro Glu Thr Glu Ala Ala Glu			
1265	1270		



<210> 18  
 <211> 1273  
 <212> PRT  
 <213> *Sinorhizobium meliloti*

<400> 18  
 Val Ser Lys Ser Ile Ile Leu Cys Arg Phe Gln Asn Gly Arg Ser Pro  
 1 5 10 15  
 Met Ser Ala Ala Asp Ala Leu Phe Gly Asn Val Ser Pro Lys Pro Asp  
 20 25 30  
 Gly Ser Glu Val Phe Arg Gln Leu Ala Gln Ala Ala Ala Glu Arg Ile  
 35 40 45  
 Leu Ile Met Asp Gly Ala Met Gly Thr Glu Ile Gln Gln Leu Gly Phe  
 50 55 60  
 Val Glu Asp His Phe Arg Gly Glu Arg Phe Gly Gly Cys Ala Cys His  
 65 70 75 80  
 Gln Gln Gly Asn Asn Asp Leu Leu Thr Leu Thr Gln Pro Lys Ala Ile  
 85 90 95  
 Glu Asp Ile His Tyr His Tyr Ala Ile Ala Gly Ala Asp Ile Leu Glu  
 100 105 110  
 Thr Asn Thr Phe Ser Ser Thr Arg Ile Ala Gln Ala Asp Tyr Gly Met  
 115 120 125  
 Glu Asp Met Val Tyr Asp Leu Asn Arg Asp Gly Ala Arg Leu Ala Arg  
 130 135 140  
 Arg Ala Ala Lys Arg Ala Glu Ala Glu Asp Gly Arg Arg Arg Phe Val  
 145 150 155 160  
 Ala Gly Ala Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp  
 165 170 175  
 Val Asn Asn Pro Gly Tyr Arg Ala Val Ser Phe Asp Asp Leu Arg Leu  
 180 185 190  
 Ala Tyr Ala Glu Gln Val Arg Gly Leu Ile Asp Gly Gly Ala Asp Ile  
 195 200 205  
 Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile  
 210 215 220  
 Phe Ala Thr Gln Glu Val Phe Ala Glu Lys Gly Val Arg Leu Pro Val  
 225 230 235 240  
 Met Ile Ser Gly Thr Ile Thr Asp Leu Ser Gly Arg Thr Leu Ser Gly  
 245 250 255  
 Gln Thr Pro Thr Ala Phe Trp Tyr Ser Val Arg His Ala Asp Pro Phe  
 260 265 270  
 Thr Ile Gly Leu Asn Cys Ala Leu Gly Ala Asn Ala Met Arg Ala His  
 275 280 285

76

Ile Asp Glu Leu Ser Ala Val Ala Asp Thr Leu Val Cys Ala Tyr Pro  
 290 295 300

Asn Ala Gly Leu Pro Asn Glu Phe Gly Arg Tyr Asp Glu Ser Pro Glu  
 305 310 315 320

Gln Met Ala Ala Gln Val Glu Gly Phe Ala Arg Asp Gly Leu Val Asn  
 325 330 335

Ile Val Gly Gly Cys Cys Gly Ser Thr Pro Ala His Ile Arg Ala Ile  
 340 345 350

Ala Glu Ala Val Ala Lys Tyr Pro Pro Arg Arg Val Pro Glu Ile Asp  
 355 360 365

Arg Arg Met Arg Leu Ser Gly Leu Glu Pro Phe Thr Leu Thr Asp Glu  
 370 375 380

Ile Pro Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala  
 385 390 395 400

Lys Phe Arg Lys Leu Ile Thr Ala Gly Asp Tyr Ala Ala Ala Leu Asp  
 405 410 415

Val Ala Arg Asp Gln Val Ala Asn Gly Ala Gln Ile Ile Asp Val Asn  
 420 425 430

Met Asp Glu Gly Leu Ile Asp Ser Lys Gln Val Met Val Glu Phe Leu  
 435 440 445

Asn Leu Val Ala Ser Glu Pro Asp Ile Ala Arg Val Pro Val Met Ile  
 450 455 460

Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys Cys Val Gln  
 465 470 475 480

Gly Lys Ala Leu Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Ala Ala  
 485 490 495

Phe Leu His His Ala Arg Leu Val Arg Ala Tyr Gly Ala Ala Val Val  
 500 505 510

Val Met Ala Phe Asp Glu Lys Gly Gln Ala Asp Thr Lys Thr Arg Lys  
 515 520 525

Val Glu Ile Cys Arg Arg Ala Tyr Arg Leu Leu Thr Glu Glu Val Gly  
 530 535 540

Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val Ala  
 545 550 555 560

Thr Gly Ile Glu Glu His Asn Asn Tyr Gly Val Asp Phe Ile Glu Ala  
 565 570 575

Thr His Glu Ile Ile Ala Ala Leu Pro His Val His Val Ser Gly Gly  
 580 585 590

Val Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg Glu  
 595 600 605

Ala Met His Ala Ile Phe Leu Tyr His Ala Ile Gln Ala Gly Met Asp  
 610 615 620

Met Gly Ile Val Asn Ala Gly Gln Leu Ala Val Tyr Asp Ala Ile Asp  
 625 630 635 640  
 Pro Glu Leu Arg Glu Thr Cys Glu Asp Val Val Leu Asn Arg Arg Ala  
 645 650 655  
 Asp Ser Thr Glu Arg Leu Leu Glu Ile Ala Glu Arg Tyr Arg Gly Lys  
 660 665 670  
 Gly Gly Ser Gln Gly Lys Glu Lys Asp Leu Ala Trp Arg Glu Trp Pro  
 675 680 685  
 Val Glu Lys Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Glu Phe  
 690 695 700  
 Ile Glu Ala Asp Thr Glu Glu Ala Arg Leu Ala Ala Glu Arg Pro Leu  
 705 710 715 720  
 His Val Ile Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp  
 725 730 735  
 Leu Phe Gly Ser Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala  
 740 745 750  
 Arg Val Met Lys Gln Ala Val Ala Val Leu Leu Pro His Met Glu Glu  
 755 760 765  
 Glu Lys Arg Ala Asn Gly Gly Gly Glu Ala Arg Glu Ser Ala Gly Lys  
 770 775 780  
 Ile Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn  
 785 790 795 800  
 Ile Val Gly Val Val Leu Ala Cys Asn Asn Tyr Glu Ile Ile Asp Leu  
 805 810 815  
 Gly Val Met Val Pro Ser Ala Lys Ile Leu Glu Val Ala Arg Glu Gln  
 820 825 830  
 Lys Val Asp Ile Val Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp  
 835 840 845  
 Glu Met Ala His Val Ala Ser Glu Leu Glu Arg Glu Gly Phe Asp Val  
 850 855 860  
 Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val  
 865 870 875 880  
 Lys Ile Asn Pro Arg Tyr Ser Leu Gly Gln Thr Val Tyr Val Thr Asp  
 885 890 895  
 Ala Ser Arg Ala Val Gly Val Val Ser Ser Leu Leu Ser Pro Glu Val  
 900 905 910  
 Arg Asp Ser Tyr Lys Lys Thr Val Arg Ala Glu Tyr Leu Lys Val Ala  
 915 920 925  
 Asp Ala His Ala Arg Asn Glu Ala Glu Lys Arg Arg Leu Pro Leu Ser  
 930 935 940  
 Gln Ala Arg Ala Asn Ala Phe Arg Ile Asp Trp Asp Ala His Gln Pro

78

945	950	955	960
Lys Val Pro Ser Phe Leu Gly Thr Arg Val Phe Glu Gly Trp Asp Leu	965	970	975
Ala Glu Leu Ala Arg Tyr Ile Asp Trp Thr Pro Phe Phe Gln Thr Trp	980	985	990
Glu Leu Lys Gly Val Phe Pro Lys Ile Leu Asp Asp Glu Arg Gln Gly	995	1000	1005
Ala Ala Ala Arg Gln Leu Phe Glu Asp Ala Gln Ala Met Val Glu Lys	1010	1015	1020
Ile Val Ala Glu Ala Trp Phe Ala Pro Lys Ala Val Ile Gly Phe Trp	1025	1030	1035
Pro Ala Ala Ser Met Gly Asp Asp Val Arg Leu Phe Ala Asp Glu Val	1045	1050	1055
Arg Glu Ala Glu Leu Ala Thr Phe Phe Thr Leu Arg Gln Gln Met Val	1060	1065	1070
Lys Arg Asp Gly Arg Pro Asn Val Ala Leu Ala Asp Phe Val Ala Pro	1075	1080	1085
Ala Ala Ser Gly Lys Arg Asp Tyr Val Gly Gly Phe Val Val Thr Ala	1090	1095	1100
Gly Ile Glu Glu Val Ala Ile Ala Glu Arg Phe Glu Arg Ala Asn Asp	1105	1110	1115
Asp Tyr Ser Ser Ile Met Val Lys Ala Leu Ala Asp Arg Phe Ala Glu	1125	1130	1135
Ala Phe Ala Glu Arg Met His Glu Tyr Val Arg Lys Glu Leu Trp Gly	1140	1145	1150
Tyr Ala Pro Asp Glu Ala Phe Thr Pro Gln Glu Leu Ile Ala Glu Pro	1155	1160	1165
Tyr Ala Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His	1170	1175	1180
Thr Glu Lys Glu Thr Leu Phe Arg Leu Leu Asp Ala Glu Ala Ala Ile	1185	1190	1195
Gly Val Arg Leu Thr Glu Ser Tyr Ala Met Trp Pro Gly Ser Ser Val	1205	1210	1215
Ser Gly Leu Tyr Val Gly His Pro Asp Ser Tyr Tyr Phe Gly Val Ala	1220	1225	1230
Lys Ile Glu Arg Asp Gln Val Glu Asp Tyr Ala Asp Arg Lys Arg Met	1235	1240	1245
Ser Val Arg Glu Val Glu Arg Trp Leu Ser Pro Ile Leu Asn Tyr Val	1250	1255	1260
Pro Met Pro Glu Thr Glu Ala Ala Glu	1265	1270	

<210> 19  
 <211> 3684  
 <212> DNA  
 <213> Escherichia coli

<220>  
 <221> CDS  
 <222> (1)..(3681)  
 <223> REC03905

<400> 19  
 gtg agc agc aaa gtg gaa caa ctg cgt gcg cag tta aat gaa cgt att 48  
 Val Ser Ser Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile  
 1 5 10 15

ctg gtg ctg gac ggc ggt atg ggc acc atg atc cag agt tat cga ctg 96  
 Leu Val Leu Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu  
 20 25 30

aac gaa gcc gat ttt cgt ggt gaa cgc ttt gcc gac tgg cca tgc gac 144  
 Asn Glu Ala Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp  
 35 40 45

ctc aaa ggc aac aac gac ctg ctg gta ctc agt aaa ccg gaa gtg atc 192  
 Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile  
 50 55 60

gcc gct atc cac aac gcc tac ttt gaa gcg ggc gcg gat atc atc gaa 240  
 Ala Ala Ile His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu  
 65 70 75 80

acc aac acc ttc aac tcc acg acc att gcg atg gcg gat tac cag atg 288  
 Thr Asn Thr Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Gln Met  
 85 90 95

gaa tcc ctg tgc gcg gaa atc aac ttt gcg gcg gcg aaa ctg gcg cga 336  
 Glu Ser Leu Ser Ala Glu Ile Asn Phe Ala Ala Ala Lys Leu Ala Arg  
 100 105 110

gct tgt gct gac gag tgg acc gcg cgc acg cca gag aaa ccg cgc tac 384  
 Ala Cys Ala Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Tyr  
 115 120 125

gtt gcc ggt gtt ctc ggc ccg acc aac cgc acg gcg tct att tct ccg 432  
 Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro  
 130 135 140

gac gtc aac gat ccg gca ttt cgt aat atc act ttt gac ggg ctg gtg 480  
 Asp Val Asn Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gly Leu Val  
 145 150 155 160

gcg gct tat cga gag tcc acc aaa gcg ctg gtg gaa ggt ggc gcg gat 528  
 Ala Ala Tyr Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Gly Ala Asp  
 165 170 175

ctg atc ctg att gaa acc gtt ttc gac acc ctt aac gcc aaa gcg gcg 576  
 Leu Ile Leu Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala  
 180 185 190

gta ttt gcg gtg aaa acg gag ttt gaa gcg ctg ggc gtt gag ctg ccg 624  
 Val Phe Ala Val Lys Thr Glu Phe Glu Ala Leu Gly Val Glu Leu Pro

80

195	200	205	
att atg atc tcc ggc acc atc acc gac gcc tcc ggg cgc acg ctc tcc Ile Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser 210 215 220			672
ggg cag acc acc gaa gca ttt tac aac tca ttg cgc cac gcc gaa gct Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala 225 230 235 240			720
ctg acc ttt ggc ctg aac tgt gcg ctg ggg ccc gat gaa ctg cgc cag Leu Thr Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln 245 250 255			768
tac gtg cag gag ctg tca cgg att gcg gaa tgc tac gtc acc gcg cac Tyr Val Gln Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His 260 265 270			816
ccg aac gcc ggg cta ccc aac gcc ttt ggt gag tac gat ctc gac gcc Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala 275 280 285			864
gac acg atg gca aaa cag ata cgt gaa tgg gcg caa gcg ggt ttt ctc Asp Thr Met Ala Lys Gln Ile Arg Glu Trp Ala Gln Ala Gly Phe Leu 290 295 300			912
aat atc gtc ggc ggc tgc tgt ggc acc acg cca caa cat att gca gcg Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Gln His Ile Ala Ala 305 310 315 320			960
atg agt cgt gca gta gaa gga tta gcg ccg cgc aaa ctg ccg gaa att Met Ser Arg Ala Val Glu Gly Leu Ala Pro Arg Lys Leu Pro Glu Ile 325 330 335			1008
ccc gta gcc tgc cgt ttg tcc ggc ctg gag ccg ctg aac att ggc gaa Pro Val Ala Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Glu 340 345 350			1056
gat agc ctg ttt gtg aac gtg ggt gaa cgc acc aac gtc acc ggt tcc Asp Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser 355 360 365			1104
gct aag ttc aag cgc ctg atc aaa gaa gag aaa tac agc gag gcg ctg Ala Lys Phe Lys Arg Leu Ile Lys Glu Glu Lys Tyr Ser Glu Ala Leu 370 375 380			1152
gat gtc gcg cgt caa cag gtg gaa aac ggc gcg cag att atc gat atc Asp Val Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile 385 390 395 400			1200
aac atg gat gaa ggg atg ctc gat gcc gaa gcg gcg atg gtg cgt ttt Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Ala Met Val Arg Phe 405 410 415			1248
ctc aat ctg att gcc ggt gaa ccg gat atc gct cgc gtg ccg att atg Leu Asn Leu Ile Ala Gly Glu Pro Asp Ile Ala Arg Val Pro Ile Met 420 425 430			1296
atc gac tcc tca aaa tgg gac gtc att gaa aaa ggt ctg aag tgt atc Ile Asp Ser Ser Lys Trp Asp Val Ile Glu Lys Gly Leu Lys Cys Ile 435 440 445			1344

81

cag ggc aaa ggc att gtt aac tct atc tcg atg aaa gag ggc gtc gat Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Met Lys Glu Gly Val Asp 450 455 460	1392
gcc ttt atc cat cac gcg aaa ttg ttg cgt cgc tac ggt gcg gca gtg Ala Phe Ile His His Ala Lys Leu Leu Arg Arg Tyr Gly Ala Ala Val 465 470 475 480	1440
gtg gta atg gcc ttt gac gaa cag gga cag gcc gat act cgc gca cgg Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Arg Ala Arg 485 490 495	1488
aaa atc gag att tgc cgt cgg gcg tac aaa atc ctc acc gaa gag gtt Lys Ile Glu Ile Cys Arg Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val 500 505 510	1536
ggc ttc ccg cca gaa gat atc atc ttc gac cca aac atc ttc gcg gtc Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val 515 520 525	1584
gca act ggc att gaa gag cac aac aac tac gcg cag gac ttt atc ggc Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Gln Asp Phe Ile Gly 530 535 540	1632
gcg tgt gaa gac atc aaa cgc gaa ctg ccg cac gcg ctg att tcc ggc Ala Cys Glu Asp Ile Lys Arg Glu Leu Pro His Ala Leu Ile Ser Gly 545 550 555 560	1680
ggc gta tct aac gtt tct ttc tcg ttc cgt ggc aac gat ccg gtg cgc Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Pro Val Arg 565 570 575	1728
gaa gcc att cac gca gtg ttc ctc tac tac gct att cgc aat ggc atg Glu Ala Ile His Ala Val Phe Leu Tyr Tyr Ala Ile Arg Asn Gly Met 580 585 590	1776
gat atg ggg atc gtc aac gcc ggg caa ctg gcg att tac gac gac cta Asp Met Gly Ile Val Asn Ala Gly Gln Leu Ala Ile Tyr Asp Asp Leu 595 600 605	1824
ccc gct gaa ctg cgc gac gcg gtg gaa gat gtg att ctt aat cgt cgc Pro Ala Glu Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg 610 615 620	1872
gac gat ggc acc gag cgt tta ctg gag ctt gcc gag aaa tat cgc ggc Asp Asp Gly Thr Glu Arg Leu Leu Glu Leu Ala Glu Lys Tyr Arg Gly 625 630 635 640	1920
agc aaa acc gac gac acc gcc aac gcc cag cag gcg gag tgg cgc tcg Ser Lys Thr Asp Asp Thr Ala Asn Ala Gln Gln Ala Glu Trp Arg Ser 645 650 655	1968
tgg gaa gtg aat aaa cgt ctg gaa tac tcg ctg gtc aaa ggc att acc Trp Glu Val Asn Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr 660 665 670	2016
gag ttt atc gag cag gat acc gaa gaa gcc cgc cag cag gct acg cgc Glu Phe Ile Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Thr Arg 675 680 685	2064
ccg att gaa gtg att gaa ggc ccg ttg atg gac ggc atg aat gtg gtc Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val	2112

690	695	700	
ggc gac ctg ttt ggc gaa ggg aaa atg ttc ctg cca cag gtg gtc aaa Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys 705' 710 715 720			2160
tcg gcg cgc gtc atg aaa cag gcg gtg gcc tac ctc gaa ccg ttt att Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile 725 730 735			2208
gaa gcc agc aaa gag cag ggc aaa acc aac ggc aag atg gtg atc gcc Glu Ala Ser Lys Glu Gln Gly Lys Thr Asn Gly Lys Met Val Ile Ala 740 745 750			2256
acc gtg aag ggc gac gtc cac gac atc ggt aaa aat atc gtt ggt gtg Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val 755 760 765			2304
gtg ctg caa tgt aac aac tac gaa att gtc gat ctc ggc gtt atg gtg Val Leu Gln Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val 770 775 780			2352
cct gcg gaa aaa att ctc cgt acc gct aaa gaa gtg aat gct gat ctg Pro Ala Glu Lys Ile Leu Arg Thr Ala Lys Glu Val Asn Ala Asp Leu 785 790 795 800			2400
att ggc ctt tcg ggg ctt atc acg ccg tcg ctg gac gag atg gtt aac Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn 805 810 815			2448
gtg gcg aaa gag atg gag cgt cag ggc ttc act att ccg tta ctg att Val Ala Lys Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile 820 825 830			2496
ggc ggc gcg acg acc tca aaa gcg cac acg gcg gtg aaa atc gag cag Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln 835 840 845			2544
aac tac agc ggc ccg acg gtg tat gtg cag aat gcc tcg cgt acc gtt Asn Tyr Ser Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val 850 855 860			2592
ggt gtg gtg gcg gcg ctg ctt tcc gat acc cag cgt gat gat ttt gtc Gly Val Val Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val 865 870 875 880			2640
gct cgt acc cgc aag gag tac gaa acc gta cgt att cag cac ggc cgc Ala Arg Thr Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Gly Arg 885 890 895			2688
aag aaa ccg cgc aca cca ccg gtc acg ctg gaa gcg gcg cgc gat aac Lys Lys Pro Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn 900 905 910			2736
gat ttc gct ttt gac tgg cag gct tac acg ccg ccg gtg gcg cac cgt Asp Phe Ala Phe Asp Trp Gln Ala Tyr Thr Pro Pro Val Ala His Arg 915 920 925			2784
ctc ggc gtg cag gaa gtc gaa gcc agc atc gaa acg ctg cgt aat tac Leu Gly Val Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr 930 935 940			2832



atc gac tgg aca ccg ttc ttt atg acc tgg tcg ctg gcc ggg aag tat	2880
Ile Asp Trp Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr	
945 950 955 960	
ccg cgc att ctg gaa gat gaa gtg gtg ggc gtt gag gcg cag cgg ctg	2928
Pro Arg Ile Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu	
965 970 975	
ttt aaa gac gcc aac gac atg ctg gat aaa tta agc gcc gag aaa acg	2976
Phe Lys Asp Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Thr	
980 985 990	
ctg aat ccg cgt ggc gtg gtg ggc ctg ttc ccg gca aac cgt gtg ggc	3024
Leu Asn Pro Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly	
995 1000 1005	
gat gac att gaa atc tac cgt gac gaa acg cgt acc cat gtg atc aac	3072
Asp Asp Ile Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Ile Asn	
1010 1015 1020	
gtc agc cac cat ctg cgt caa cag acc gaa aaa aca ggc ttc gct aac	3120
Val Ser His His Leu Arg Gln Gln Thr Glu Lys Thr Gly Phe Ala Asn	
1025 1030 1035 1040	
tac tgt ctc gct gac ttc gtt gcg ccg aag ctt tct ggt aaa gca gat	3168
Tyr Cys Leu Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp	
1045 1050 1055	
tac atc ggc gca ttt gcc gtg act ggc ggg ctg gaa gag gac gca ctg	3216
Tyr Ile Gly Ala Phe Ala Val Thr Gly Gly Leu Glu Glu Ala Leu	
1060 1065 1070	
gct gat gcc ttt gaa gcg cag cac gat gat tac aac aaa atc atg gtg	3264
Ala Asp Ala Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val	
1075 1080 1085	
aaa gcg ctt gcc gac cgt tta gcc gaa gcc ttt gcg gag tat ctc cat	3312
Lys Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr Leu His	
1090 1095 1100	
gag cgt gtg cgt aaa gtc tac tgg ggc tat gcg ccg aac gag aac ctc	3360
Glu Arg Val Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Asn Leu	
1105 1110 1115 1120	
agc aac gaa gag ctg atc cgc gaa aac tac cag ggc atc cgt ccg gca	3408
Ser Asn Glu Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala	
1125 1130 1135	
ccg gcc tat ccg gcc tgc ccg gaa cat acg gaa aaa gcc acc atc tgg	3456
Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Ala Thr Ile Trp	
1140 1145 1150	
gag ctg ctg gaa gtg gaa aaa cac act ggc atg aaa ctc aca gaa tct	3504
Glu Leu Leu Glu Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser	
1155 1160 1165	
ttc gcc atg tgg ccc ggt gca tgc gtt tgc ggt tgg tac ttc agc cac	3552
Phe Ala Met Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His	
1170 1175 1180	
ccg gac agc aag tac tac gct gta gca caa att cag cgc gat cag gtt	3600
Pro Asp Ser Lys Tyr Tyr Ala Val Ala Gln Ile Gln Arg Asp Gln Val	

84

1185                      1190                      1195                      1200  
 gaa gat tat gcc cgc cgt aaa ggt atg agc gtt acc gaa gtt gag cgc 3648  
 Glu Asp Tyr Ala Arg Arg Lys Gly Met Ser Val Thr Glu Val Glu Arg  
                     1205                      1210                      1215

tgg ctg gca ccg aat ctg ggg tat gac gcg gac tga 3684  
 Trp Leu Ala Pro Asn Leu Gly Tyr Asp Ala Asp  
                     1220                      1225

<210> 20  
 <211> 1227  
 <212> PRT  
 <213> Escherichia coli

<400> 20  
 Val Ser Ser Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile  
   1                    5                    10                    15

Leu Val Leu Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu  
                     20                    25                    30

Asn Glu Ala Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp  
                     35                    40                    45

Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile  
                     50                    55                    60

Ala Ala Ile His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu  
                     65                    70                    75                    80

Thr Asn Thr Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Gln Met  
                     85                    90                    95

Glu Ser Leu Ser Ala Glu Ile Asn Phe Ala Ala Ala Lys Leu Ala Arg  
                     100                    105                    110

Ala Cys Ala Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Tyr  
                     115                    120                    125

Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro  
                     130                    135                    140

Asp Val Asn Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gly Leu Val  
                     145                    150                    155                    160

Ala Ala Tyr Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Gly Ala Asp  
                     165                    170                    175

Leu Ile Leu Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala  
                     180                    185                    190

Val Phe Ala Val Lys Thr Glu Phe Glu Ala Leu Gly Val Glu Leu Pro  
                     195                    200                    205

Ile Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser  
                     210                    215                    220

Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala  
                     225                    230                    235                    240

85

Leu Thr Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln  
 245 250 255  
 Tyr Val Gln Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His  
 260 265 270  
 Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala  
 275 280 285  
 Asp Thr Met Ala Lys Gln Ile Arg Glu Trp Ala Gln Ala Gly Phe Leu  
 290 295 300  
 Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Gln His Ile Ala Ala  
 305 310 315 320  
 Met Ser Arg Ala Val Glu Gly Leu Ala Pro Arg Lys Leu Pro Glu Ile  
 325 330 335  
 Pro Val Ala Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Glu  
 340 345 350  
 Asp Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser  
 355 360 365  
 Ala Lys Phe Lys Arg Leu Ile Lys Glu Glu Lys Tyr Ser Glu Ala Leu  
 370 375 380  
 Asp Val Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile  
 385 390 395 400  
 Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Ala Met Val Arg Phe  
 405 410 415  
 Leu Asn Leu Ile Ala Gly Glu Pro Asp Ile Ala Arg Val Pro Ile Met  
 420 425 430  
 Ile Asp Ser Ser Lys Trp Asp Val Ile Glu Lys Gly Leu Lys Cys Ile  
 435 440 445  
 Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Met Lys Glu Gly Val Asp  
 450 455 460  
 Ala Phe Ile His His Ala Lys Leu Leu Arg Arg Tyr Gly Ala Ala Val  
 465 470 475 480  
 Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Arg Ala Arg  
 485 490 495  
 Lys Ile Glu Ile Cys Arg Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val  
 500 505 510  
 Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val  
 515 520 525  
 Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Gln Asp Phe Ile Gly  
 530 535 540  
 Ala Cys Glu Asp Ile Lys Arg Glu Leu Pro His Ala Leu Ile Ser Gly  
 545 550 555 560  
 Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Pro Val Arg  
 565 570 575

Glu Ala Ile His Ala Val Phe Leu Tyr Tyr Ala Ile Arg Asn Gly Met  
                   580                                  585                                  590

Asp Met Gly Ile Val Asn Ala Gly Gln Leu Ala Ile Tyr Asp Asp Leu  
                   595                                  600                                  605

Pro Ala Glu Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg  
                   610                                  615                                  620

Asp Asp Gly Thr Glu Arg Leu Leu Glu Leu Ala Glu Lys Tyr Arg Gly  
                   625                                  630                                  635                                  640

Ser Lys Thr Asp Asp Thr Ala Asn Ala Gln Gln Ala Glu Trp Arg Ser  
                                   645                                  650                                  655

Trp Glu Val Asn Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr  
                                   660                                  665                                  670

Glu Phe Ile Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Thr Arg  
                   675                                  680                                  685

Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val  
                   690                                  695                                  700

Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys  
                   705                                  710                                  715                                  720

Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile  
                                   725                                  730                                  735

Glu Ala Ser Lys Glu Gln Gly Lys Thr Asn Gly Lys Met Val Ile Ala  
                                   740                                  745                                  750

Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val  
                                   755                                  760                                  765

Val Leu Gln Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val  
                   770                                  775                                  780

Pro Ala Glu Lys Ile Leu Arg Thr Ala Lys Glu Val Asn Ala Asp Leu  
                   785                                  790                                  795                                  800

Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn  
                                   805                                  810                                  815

Val Ala Lys Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile  
                                   820                                  825                                  830

Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln  
                   835                                  840                                  845

Asn Tyr Ser Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val  
                   850                                  855                                  860

Gly Val Val Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val  
                   865                                  870                                  875                                  880

Ala Arg Thr Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Gly Arg  
                                   885                                  890                                  895

Lys Lys Pro Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn

87

900	905	910
Asp Phe Ala Phe Asp Trp Gln Ala Tyr Thr Pro Pro Val Ala His Arg 915 920 925		
Leu Gly Val Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr 930 935 940		
Ile Asp Trp Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr 945 950 955 960		
Pro Arg Ile Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu 965 970 975		
Phe Lys Asp Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Thr 980 985 990		
Leu Asn Pro Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly 995 1000 1005		
Asp Asp Ile Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Ile Asn 1010 1015 1020		
Val Ser His His Leu Arg Gln Gln Thr Glu Lys Thr Gly Phe Ala Asn 1025 1030 1035 1040		
Tyr Cys Leu Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp 1045 1050 1055		
Tyr Ile Gly Ala Phe Ala Val Thr Gly Gly Leu Glu Glu Asp Ala Leu 1060 1065 1070		
Ala Asp Ala Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val 1075 1080 1085		
Lys Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr Leu His 1090 1095 1100		
Glu Arg Val Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Asn Leu 1105 1110 1115 1120		
Ser Asn Glu Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala 1125 1130 1135		
Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Ala Thr Ile Trp 1140 1145 1150		
Glu Leu Leu Glu Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser 1155 1160 1165		
Phe Ala Met Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His 1170 1175 1180		
Pro Asp Ser Lys Tyr Tyr Ala Val Ala Gln Ile Gln Arg Asp Gln Val 1185 1190 1195 1200		
Glu Asp Tyr Ala Arg Arg Lys Gly Met Ser Val Thr Glu Val Glu Arg 1205 1210 1215		
Trp Leu Ala Pro Asn Leu Gly Tyr Asp Ala Asp 1220 1225		

<210> 21  
 <211> 3771  
 <212> DNA  
 <213> Salmonella typhimurium

<220>  
 <221> CDS  
 <222> (1) .. (3768)  
 <223> RSY02996

<400> 21  
 atg tct cat gtt gcc cgt tgt tct ctt ttc cgc cag cac gct ttg tgc 48  
 Met Ser His Val Ala Arg Cys Ser Leu Phe Arg Gln His Ala Leu Cys  
 1 5 10 15

cag tat ggc tcg tta cgt gga gcg ttg tcg gga gcg agt gtg agc agc 96  
 Gln Tyr Gly Ser Leu Arg Gly Ala Leu Ser Gly Ala Ser Val Ser Ser  
 20 25 30

aaa gtt gaa caa ctg cgt gcg cag tta aat gaa cgt att ctg gtg ctg 144  
 Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile Leu Val Leu  
 35 40 45

gac ggc ggt atg ggc act atg atc cag agc tat cgt cta cat gaa gaa 192  
 Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu His Glu Glu  
 50 55 60

gat ttc cgc ggg gag cgc ttt gcc gac tgg ccc tgc gac ctg aaa ggc 240  
 Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp Leu Lys Gly  
 65 70 75 80

aac aat gac ctg ctg gtc ctc agc aag ccg gag gtg atc gcc gct atc 288  
 Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile Ala Ala Ile  
 85 90 95

cac aac gcc tac ttt gag gct ggc gcg gat atc atc gaa acc aac acc 336  
 His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr  
 100 105 110

ttt aac tcg aca acc att gcg atg gcg gat tac cgg atg gaa tcc ctg 384  
 Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Arg Met Glu Ser Leu  
 115 120 125

tcg gcg gaa att aac tat gcg gcg gcc aaa ctg gcg cgc gcc tgc gcc 432  
 Ser Ala Glu Ile Asn Tyr Ala Ala Ala Lys Leu Ala Arg Ala Cys Ala  
 130 135 140

gat gaa tgg acg gcg cga aca cca gaa aaa cca cgc ttt gtt gcg ggc 480  
 Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Phe Val Ala Gly  
 145 150 155 160

gtg ctt ggt cca act aac cgc acg gcc tcc att tcg ccg gac gtc aac 528  
 Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp Val Asn  
 165 170 175

gac ccg gcg ttt cgt aat atc acc ttc gat cag ctg gtg gcg gcc tac 576  
 Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gln Leu Val Ala Ala Tyr  
 180 185 190

cgt gaa tcc acc aaa gcg ctg gtg gaa ggt ggc gca gat ctg att ctg 624  
 Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Gly Ala Asp Leu Ile Leu

195	200	205	
att gag acc gtt ttt gac acc ctg aat gcg aaa gcg gcg gtg ttt gcg Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Val Phe Ala 210 215 220			672
gtg aaa gaa gag ttt gaa gcg ctg ggc gtt gac ctg ccg atc atg att Val Lys Glu Glu Phe Glu Ala Leu Gly Val Asp Leu Pro Ile Met Ile 225 230 235 240			720
tcc ggc acc atc acc gac gcc tct ggc cgt acg ctt tcc ggc cag act Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr 245 250 255			768
acc gaa gcc ttt tat aac tcg ctg cgc cac gcc gag gcg ctc act ttt Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala Leu Thr Phe 260 265 270			816
ggc ctt aac tgc gca ctg ggg cca gat gaa ctg cgc cag tac gtc cag Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln Tyr Val Gln 275 280 285			864
gaa ctg tcg cgg att gcc gaa tgc tac gtc acc gcg cac ccg aac gcc Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His Pro Asn Ala 290 295 300			912
ggc ctg ccg aac gct ttc ggc gag tat gac ctc gac gcc gac acc atg Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala Asp Thr Met 305 310 315 320			960
gcg aaa cag att cgc gaa tgg gcg gaa gcg ggc ttc ctg aat atc gtt Ala Lys Gln Ile Arg Glu Trp Ala Glu Ala Gly Phe Leu Asn Ile Val 325 330 335			1008
ggc ggc tgc tgc ggc acc acg ccg gag cat att gcg gcg atg agc cgc Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala Ala Met Ser Arg 340 345 350			1056
gcc gtt gcc ggt ttg ctg ccg cgc cag ctg ccg gat atc ccg gtc gcc Ala Val Ala Gly Leu Leu Pro Arg Gln Leu Pro Asp Ile Pro Val Ala 355 360 365			1104
tgc cgc ctt tcc ggc ctg gag ccg ctg aac att ggc gac gat agc ctg Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Asp Asp Ser Leu 370 375 380			1152
ttt gtg aac gtc ggc gaa cgt act aac gtc acc ggc tcg gcc aaa ttt Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Lys Phe 385 390 395 400			1200
aaa cgg ctg atc aaa gaa gag aaa tac agc gaa gcg ctg gat gtc gcc Lys Arg Leu Ile Lys Glu Glu Lys Tyr Ser Glu Ala Leu Asp Val Ala 405 410 415			1248
cgt cag cag gta gaa agc ggc gcg cag att att gat atc aat atg gat Arg Gln Gln Val Glu Ser Gly Ala Gln Ile Ile Asp Ile Asn Met Asp 420 425 430			1296
gag ggg atg ctc gac gcc gaa gcg gcg atg gtg cgt ttc ctc agc ctg Glu Gly Met Leu Asp Ala Glu Ala Ala Met Val Arg Phe Leu Ser Leu 435 440 445			1344

90

att gcc ggt gag ccg gac	att gcc cgt gta cca atc atg atc gac tcc	1392
Ile Ala Gly Glu Pro Asp	Ile Ala Arg Val Pro Ile Met Ile Asp Ser	
450	455 460	
tcc aaa tgg gag gtt atc gaa aaa ggg ctg aag tgc att cag ggt aaa	1440	
Ser Lys Trp Glu Val Ile Glu Lys Gly Leu Lys Cys Ile Gln Gly Lys		
465	470 475 480	
ggc atc gtc aac tct att tcg atg aaa gag ggc gtg gaa gcc ttt att	1488	
Gly Ile Val Asn Ser Ile Ser Met Lys Glu Gly Val Glu Ala Phe Ile		
485	490 495	
cat cat gcg aag ctt ctg cgt cgc tac ggc gcg gca gtg gtg gtg atg	1536	
His His Ala Lys Leu Leu Arg Arg Tyr Gly Ala Ala Val Val Val Met		
500	505 510	
gcc ttt gat gag cag ggg cag gcc gat acc cgc gcg cgt aaa atc gaa	1584	
Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Arg Ala Arg Lys Ile Glu		
515	520 525	
att tgc cgc cga gcc tac aaa att ctc acc gaa gag gtg ggt ttc ccg	1632	
Ile Cys Arg Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val Gly Phe Pro		
530	535 540	
ccg gaa gac atc atc ttc gac ccg aat atc ttc gcc gtg gcg acc ggt	1680	
Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val Ala Thr Gly		
545	550 555 560	
att gaa gag cac aac aac tac gcg cag gac ttt atc ggc gct tgc gaa	1728	
Ile Glu Glu His Asn Asn Tyr Ala Gln Asp Phe Ile Gly Ala Cys Glu		
565	570 575	
gac atc aaa cgc gag ctg ccg cac gcg ctg atc tcc ggc ggc gtg tct	1776	
Asp Ile Lys Arg Glu Leu Pro His Ala Leu Ile Ser Gly Gly Val Ser		
580	585 590	
aac gtg tcc ttc tcg ttc cgc ggc aac gac ccg gta cgt gag gct atc	1824	
Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Pro Val Arg Glu Ala Ile		
595	600 605	
cac gcg gta ttc ctc tac tac gcc atc cgc aac ggt atg gac atg ggc	1872	
His Ala Val Phe Leu Tyr Tyr Ala Ile Arg Asn Gly Met Asp Met Gly		
610	615 620	
atc gtc aac gcc gga cag ctg gct atc tac gac gac ctg ccc gcc gag	1920	
Ile Val Asn Ala Gly Gln Leu Ala Ile Tyr Asp Asp Leu Pro Ala Glu		
625	630 635 640	
ctg cgc gat gcg gtt gaa gat gtc att ctt aac cgt cgc gat gac ggc	1968	
Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg Asp Asp Gly		
645	650 655	
act gag cgt ttg ctg gat ttg gcg gag aaa tac cgc ggc agc aaa acc	2016	
Thr Glu Arg Leu Leu Asp Leu Ala Glu Lys Tyr Arg Gly Ser Lys Thr		
660	665 670	
gac gaa gct gcc aac gcc cag cag gcg gaa tgg cgt agc tgg gac gtg	2064	
Asp Glu Ala Ala Asn Ala Gln Gln Ala Glu Trp Arg Ser Trp Asp Val		
675	680 685	
aaa aag cgt ctc gaa tac tcg ctg gtg aaa ggc att acc gaa ttt atc	2112	
Lys Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr Glu Phe Ile		



91

690	695	700	
gaa cag gat acc gaa gaa gcc cgt cag cag gcc gcc cgc ccg att gag Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Ile Glu 705 710 715 720			2160
gtg att gaa ggg ccg ttg atg gac ggc atg aac gtg gtc ggc gac ctg Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu 725 730 735			2208
ttc ggc gaa ggg aaa atg ttc ctg ccg cag gtg gtg aaa tcc gct cgc Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg 740 745 750			2256
gtg atg aaa caa gcg gtg gcc tac ctg gag ccg ttt att gaa gcc agc Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile Glu Ala Ser 755 760 765			2304
aaa gaa aaa ggc tcc agc aac ggc aag atg gtg atc gcc acc gtg aag Lys Glu Lys Gly Ser Ser Asn Gly Lys Met Val Ile Ala Thr Val Lys 770 775 780			2352
ggc gat gtg cac gat att ggt aaa aat atc gtt ggc gtg gtg ctg caa Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln 785 790 795 800			2400
tgt aac aac tac gaa atc gtc gat ctt ggc gtg atg gtg cca gcg gag Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val Pro Ala Glu 805 810 815			2448
aaa atc ctc aga acg gcg cgt gaa gtg aat gcc gat ctg atc ggt ctt Lys Ile Leu Arg Thr Ala Arg Glu Val Asn Ala Asp Leu Ile Gly Leu 820 825 830			2496
tcc ggg ctg att acc ccg tcc ctg gac gaa atg gtc aac gtg gcg aaa Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn Val Ala Lys 835 840 845			2544
gag atg gag cgt cag ggc ttt act atc ccg cta ctg atc ggc ggc gca Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile Gly Gly Ala 850 855 860			2592
acc act tcc aaa gcg cat acg gcg gtg aaa atc gag cag aac tac agc Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln Asn Tyr Ser 865 870 875 880			2640
ggc ccg acg gtc tac gtg cag aat gct tcc cgt acc gtg ggc gtg gtg Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val Gly Val Val 885 890 895			2688
gcg gcg cta ctc tcc gac acc cag cgt gat gac ttt gtc gcc cgt acc Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val Ala Arg Thr 900 905 910			2736
cgc aaa gag tac gaa acc gtg cgt att cag cac gcc cgc aaa aaa ccg Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Ala Arg Lys Lys Pro 915 920 925			2784
cgc acg ccg ccg gtc acg ctg gag gcg gcg cgc gat aac gat ctg gca Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn Asp Leu Ala 930 935 940			2832

92

ttc gat tgg gaa cgc tac acc ccg ccg gta gcc cac cgt ctg ggc gtg	2880
Phe Asp Trp Glu Arg Tyr Thr Pro Pro Val Ala His Arg Leu Gly Val	
945 950 955 960	
cag gag gtg gaa gcc agc atc gaa act ctg cgc aac tat atc gac tgg	2928
Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr Ile Asp Trp	
965 970 975	
acg ccg ttc ttt atg acc tgg tgg ctg gcc ggc aaa tac ccg cgc att	2976
Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr Pro Arg Ile	
980 985 990	
ctg gaa gat gag gtg gtg ggc gtt gag gcg cag cgt ctg ttt aaa gac	3024
Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu Phe Lys Asp	
995 1000 1005	
gcc aat gat atg ctg gat aaa ctg agc gcc gag aaa ctg ttg aat ccg	3072
Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Leu Leu Asn Pro	
1010 1015 1020	
cgt ggc gtg gtg ggc ctg ttc ccg gcg aac cgt gtg ggt gac gac atc	3120
Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly Asp Asp Ile	
1025 1030 1035 1040	
gaa atc tat cgc gac gaa acc cgt act cat gtt ctg acg gtc agc cac	3168
Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Leu Thr Val Ser His	
1045 1050 1055	
cac ctg cgc cag cag acc gag aaa gtc ggt ttt gcc aac tac tgt ctg	3216
His Leu Arg Gln Gln Thr Glu Lys Val Gly Phe Ala Asn Tyr Cys Leu	
1060 1065 1070	
gcg gat ttt gtc gcg ccg aaa ctg agc gcc aaa gcg gat tac atc ggt	3264
Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp Tyr Ile Gly	
1075 1080 1085	
gct ttc gcg gtg acc ggc ggt ctg gag gag gat gcg ctg gcg gac gcc	3312
Ala Phe Ala Val Thr Gly Gly Leu Glu Glu Asp Ala Leu Ala Asp Ala	
1090 1095 1100	
ttc gaa gcg caa cac gac gac tat aac aag atc atg gtg aaa gcg att	3360
Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val Lys Ala Ile	
1105 1110 1115 1120	
gcc gac cgt ctg gcg gaa gcg ttc gcg gaa tat ctg cat gag cgt gta	3408
Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr Leu His Glu Arg Val	
1125 1130 1135	
cgt aag gtt tac tgg gga tat gcg ccg aac gag agc ctg agt aac gac	3456
Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Ser Leu Ser Asn Asp	
1140 1145 1150	
gaa tta atc cgc gaa aac tac cag ggg att cgc ccg gcg ccg ggt tat	3504
Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr	
1155 1160 1165	
cct gcc tgc ccg gaa cat acc gaa aaa ggc act atc tgg cag cta ctg	3552
Pro Ala Cys Pro Glu His Thr Glu Lys Gly Thr Ile Trp Gln Leu Leu	
1170 1175 1180	
gat gtc gaa aaa cac acc ggg atg aag ctc acc gaa tct ttc gcc atg	3600
Asp Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser Phe Ala Met	

93

1185                      1190                      1195                      1200

tgg cca ggc gcg tcg gtc tcc ggc tgg tac ttc agc cat cct gag agc    3648  
 Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His Pro Glu Ser  
                          1205                      1210                      1215

aaa tac ttc gcg gta gcg cag atc caa cgc gat cag gtg aca gat tat    3696  
 Lys Tyr Phe Ala Val Ala Gln Ile Gln Arg Asp Gln Val Thr Asp Tyr  
                          1220                      1225                      1230

gct ttc cgt aaa gga atg agc gtt gag gat gtt gag cgg tgg ctc gcg    3744  
 Ala Phe Arg Lys Gly Met Ser Val Glu Asp Val Glu Arg Trp Leu Ala  
                          1235                      1240                      1245

ccg aac ctg ggt tac gat gcg gac tga                                      3771  
 Pro Asn Leu Gly Tyr Asp Ala Asp  
                          1250                      1255

<210> 22  
 <211> 1256  
 <212> PRT  
 <213> Salmonella typhimurium

<400> 22  
 Met Ser His Val Ala Arg Cys Ser Leu Phe Arg Gln His Ala Leu Cys  
   1                          5                          10                          15

Gln Tyr Gly Ser Leu Arg Gly Ala Leu Ser Gly Ala Ser Val Ser Ser  
                           20                          25                          30

Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile Leu Val Leu  
                           35                          40                          45

Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu His Glu Glu  
                           50                          55                          60

Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp Leu Lys Gly  
                           65                          70                          75                          80

Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile Ala Ala Ile  
                           85                          90                          95

His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr  
                           100                          105                          110

Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Arg Met Glu Ser Leu  
                           115                          120                          125

Ser Ala Glu Ile Asn Tyr Ala Ala Ala Lys Leu Ala Arg Ala Cys Ala  
                           130                          135                          140

Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Phe Val Ala Gly  
                           145                          150                          155                          160

Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp Val Asn  
                           165                          170                          175

Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gln Leu Val Ala Ala Tyr  
                           180                          185                          190

Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Gly Ala Asp Leu Ile Leu

195	200	205
Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Val Phe Ala 210 215 220		
Val Lys Glu Glu Phe Glu Ala Leu Gly Val Asp Leu Pro Ile Met Ile 225 230 235 240		
Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr 245 250 255		
Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala Leu Thr Phe 260 265 270		
Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln Tyr Val Gln 275 280 285		
Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His Pro Asn Ala 290 295 300		
Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala Asp Thr Met 305 310 315 320		
Ala Lys Gln Ile Arg Glu Trp Ala Glu Ala Gly Phe Leu Asn Ile Val 325 330 335		
Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala Ala Met Ser Arg 340 345 350		
Ala Val Ala Gly Leu Leu Pro Arg Gln Leu Pro Asp Ile Pro Val Ala 355 360 365		
Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Asp Asp Ser Leu 370 375 380		
Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Lys Phe 385 390 395 400		
Lys Arg Leu Ile Lys Glu Glu Lys Tyr Ser Glu Ala Leu Asp Val Ala 405 410 415		
Arg Gln Gln Val Glu Ser Gly Ala Gln Ile Ile Asp Ile Asn Met Asp 420 425 430		
Glu Gly Met Leu Asp Ala Glu Ala Ala Met Val Arg Phe Leu Ser Leu 435 440 445		
Ile Ala Gly Glu Pro Asp Ile Ala Arg Val Pro Ile Met Ile Asp Ser 450 455 460		
Ser Lys Trp Glu Val Ile Glu Lys Gly Leu Lys Cys Ile Gln Gly Lys 465 470 475 480		
Gly Ile Val Asn Ser Ile Ser Met Lys Glu Gly Val Glu Ala Phe Ile 485 490 495		
His His Ala Lys Leu Leu Arg Arg Tyr Gly Ala Ala Val Val Met 500 505 510		
Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Arg Ala Arg Lys Ile Glu 515 520 525		

95

Ile Cys Arg Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val Gly Phe Pro  
 530 535 540

Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val Ala Thr Gly  
 545 550 555 560

Ile Glu Glu His Asn Asn Tyr Ala Gln Asp Phe Ile Gly Ala Cys Glu  
 565 570 575

Asp Ile Lys Arg Glu Leu Pro His Ala Leu Ile Ser Gly Gly Val Ser  
 580 585 590

Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Pro Val Arg Glu Ala Ile  
 595 600 605

His Ala Val Phe Leu Tyr Tyr Ala Ile Arg Asn Gly Met Asp Met Gly  
 610 615 620

Ile Val Asn Ala Gly Gln Leu Ala Ile Tyr Asp Asp Leu Pro Ala Glu  
 625 630 635 640

Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg Asp Asp Gly  
 645 650 655

Thr Glu Arg Leu Leu Asp Leu Ala Glu Lys Tyr Arg Gly Ser Lys Thr  
 660 665 670

Asp Glu Ala Ala Asn Ala Gln Gln Ala Glu Trp Arg Ser Trp Asp Val  
 675 680 685

Lys Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr Glu Phe Ile  
 690 695 700

Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Ile Glu  
 705 710 715 720

Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu  
 725 730 735

Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg  
 740 745 750

Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile Glu Ala Ser  
 755 760 765

Lys Glu Lys Gly Ser Ser Asn Gly Lys Met Val Ile Ala Thr Val Lys  
 770 775 780

Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln  
 785 790 795 800

Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val Pro Ala Glu  
 805 810 815

Lys Ile Leu Arg Thr Ala Arg Glu Val Asn Ala Asp Leu Ile Gly Leu  
 820 825 830

Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn Val Ala Lys  
 835 840 845

Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile Gly Gly Ala  
 850 855 860

Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln Asn Tyr Ser  
 865 870 875 880  
 Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val Gly Val Val  
 885 890 895  
 Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val Ala Arg Thr  
 900 905 910  
 Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Ala Arg Lys Lys Pro  
 915 920 925  
 Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn Asp Leu Ala  
 930 935 940  
 Phe Asp Trp Glu Arg Tyr Thr Pro Pro Val Ala His Arg Leu Gly Val  
 945 950 955 960  
 Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr Ile Asp Trp  
 965 970 975  
 Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr Pro Arg Ile  
 980 985 990  
 Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu Phe Lys Asp  
 995 1000 1005  
 Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Leu Leu Asn Pro  
 1010 1015 1020  
 Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly Asp Asp Ile  
 1025 1030 1035 1040  
 Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Leu Thr Val Ser His  
 1045 1050 1055  
 His Leu Arg Gln Gln Thr Glu Lys Val Gly Phe Ala Asn Tyr Cys Leu  
 1060 1065 1070  
 Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp Tyr Ile Gly  
 1075 1080 1085  
 Ala Phe Ala Val Thr Gly Gly Leu Glu Glu Asp Ala Leu Ala Asp Ala  
 1090 1095 1100  
 Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val Lys Ala Ile  
 1105 1110 1115 1120  
 Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr Leu His Glu Arg Val  
 1125 1130 1135  
 Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Ser Leu Ser Asn Asp  
 1140 1145 1150  
 Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr  
 1155 1160 1165  
 Pro Ala Cys Pro Glu His Thr Glu Lys Gly Thr Ile Trp Gln Leu Leu  
 1170 1175 1180  
 Asp Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser Phe Ala Met

97

1185                      1190                      1195                      1200

Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His Pro Glu Ser  
1205                      1210                      1215

Lys Tyr Phe Ala Val Ala Gln Ile Gln Arg Asp Gln Val Thr Asp Tyr  
1220                      1225                      1230

Ala Phe Arg Lys Gly Met Ser Val Glu Asp Val Glu Arg Trp Leu Ala  
1235                      1240                      1245

Pro Asn Leu Gly Tyr Asp Ala Asp  
1250                      1255

&lt;210&gt; 23

&lt;211&gt; 3771

&lt;212&gt; DNA

&lt;213&gt; Salmonella typhi

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(3768)

&lt;223&gt; RTY03686

&lt;400&gt; 23

atg tct cat gtt gcc cgt tgt tct ctt ttc cgc cag cac gct ttg tgc 48  
Met Ser His Val Ala Arg Cys Ser Leu Phe Arg Gln His Ala Leu Cys  
1                      5                      10                      15

cag tat ggc tcg tta cgt gga gcg ttg tcg gga gcg agt gtg agc agc 96  
Gln Tyr Gly Ser Leu Arg Gly Ala Leu Ser Gly Ala Ser Val Ser Ser  
20                      25                      30

aaa gtt gaa caa ctg cgt gcg cag tta aat gaa cgt att ctg gtg ctg 144  
Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile Leu Val Leu  
35                      40                      45

gac ggc ggt atg ggc acc atg atc cag agc tat cgt cta cat gaa gaa 192  
Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu His Glu Glu  
50                      55                      60

gat ttc cgc ggg gag cgc ttt gcc gac tgg ccc tgc gac ctg aaa ggc 240  
Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp Leu Lys Gly  
65                      70                      75                      80

aac aat gac ctg ctg gtc ctc agc aag ccg gag gtg atc gcc gct atc 288  
Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile Ala Ala Ile  
85                      90                      95

cac aac gcc tac ttt gag gct ggc gcg gat atc atc gaa acc aac acc 336  
His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr  
100                      105                      110

ttt aac tcg aca acc att gcg atg gcg gat tac cgg atg gaa tcc ctg 384  
Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Arg Met Glu Ser Leu  
115                      120                      125

tcg gcg gaa att aac tat gcg gcg gcc aaa ctg gcg cgc gcc tgc gcc 432  
Ser Ala Glu Ile Asn Tyr Ala Ala Ala Lys Leu Ala Arg Ala Cys Ala  
130                      135                      140

98

gat gaa tgg acg gcg cga aca cca gaa aaa cca cgc ttt gtt gcg ggc Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Phe Val Ala Gly 145 150 155 160	480
gtg ctt ggt cca act aac cgc acg gcc tcc att tcg ccg gac gtc aac Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp Val Asn 165 170 175	528
gac ccg gcg ttt cgt aat atc acc ttc gat cag ctg gtg gcg gcc tac Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gln Leu Val Ala Ala Tyr 180 185 190	576
cgt gaa tcc acc aaa gcg ctg gtg gaa ggc ggg gcg gac ctg atc ctg Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Ala Asp Leu Ile Leu 195 200 205	624
att gaa act gtc ttc gac acc ctc aac gcc aaa gcg gcg gtg ttt gcg Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Val Phe Ala 210 215 220	672
gtg aaa gaa gag ttt gaa gcg ctg ggc gtt gat ctg ccg atc atg att Val Lys Glu Glu Phe Glu Ala Leu Gly Val Asp Leu Pro Ile Met Ile 225 230 235 240	720
tcc ggc acc atc acc gac gcc tct ggc cgt acg ctt tcc ggc cag acg Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr 245 250 255	768
acc gaa gcc ttt tat aac tcg ctg cgc cac gcc gag gcg ctc act ttt Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala Leu Thr Phe 260 265 270	816
ggc ctt aac tgc gcg ctg ggg cca gat gaa ctg cgc cag tac gtc cag Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln Tyr Val Gln 275 280 285	864
gaa ctg tcg ccg att gcc gaa tgc tac gtc acc gcg cac ccg aac gcc Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His Pro Asn Ala 290 295 300	912
ggc ctg ccg aac gct ttc ggc gag tac gac ctc gac gcc gac acc atg Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala Asp Thr Met 305 310 315 320	960
gcg aaa cag att cgc gaa tgg gcg gaa gcg ggc ttc ctg aat atc gtt Ala Lys Gln Ile Arg Glu Trp Ala Glu Ala Gly Phe Leu Asn Ile Val 325 330 335	1008
ggc ggc tgc tgc ggc acc acg ccg gag cat att gcg gcg atg agc cgc Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala Ala Met Ser Arg 340 345 350	1056
gcc gtt gcc ggt ttg tcg ccg cgc cag ctg ccg gat atc ccg gtg gcc Ala Val Ala Gly Leu Ser Pro Arg Gln Leu Pro Asp Ile Pro Val Ala 355 360 365	1104
tgc cgc ctt tcc ggc ctg gag ccg ctg aac att ggt gac gat agc ctg Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Asp Asp Ser Leu 370 375 380	1152
ttt gtc aac gtc ggc gaa cgt act aac gtc acc ggc tcg gcc aaa ttt Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Lys Phe	1200



99

385	390	395	400	
aaa cgc ttg atc	aaa gaa gag aaa tac agc gaa gcg ctg gat gtc gcc	1248		
Lys Arg Leu Ile	Lys Glu Glu Lys Tyr Ser Glu Ala Leu Asp Val Ala			
	405 410 415			
cgt cag cag gtc	gaa agc ggc gcg cag att att gat atc aat atg gat	1296		
Arg Gln Gln Val	Glu Ser Gly Ala Gln Ile Ile Asp Ile Asn Met Asp			
	420 425 430			
gag ggg atg ctc	gac gcc gaa gcg gcg atg gtg cgt ttc ctc agc ctg	1344		
Glu Gly Met Leu	Asp Ala Glu Ala Ala Met Val Arg Phe Leu Ser Leu			
	435 440 445			
att gcc ggt gag	ccg gac att gcc cgt gta cca atc atg att gac tcc	1392		
Ile Ala Gly Glu	Pro Asp Ile Ala Arg Val Pro Ile Met Ile Asp Ser			
	450 455 460			
tcc aaa tgg gag	gtt atc gaa aaa ggg ctg aag tgc att cag ggt aaa	1440		
Ser Lys Trp Glu	Val Ile Glu Lys Gly Leu Lys Cys Ile Gln Gly Lys			
	465 470 475 480			
ggc atc gtc aac	tct att tcg atg aaa gag ggc gtg gaa gcc ttt att	1488		
Gly Ile Val Asn	Ser Ile Ser Met Lys Glu Gly Val Glu Ala Phe Ile			
	485 490 495			
cat cat gcg aag	ttg cta cgt cgc tac ggc gcc gca gtg gtg gtg atg	1536		
His His Ala Lys	Leu Leu Arg Arg Tyr Gly Ala Ala Val Val Val Met			
	500 505 510			
gct ttt gat gag	cag ggg cag gcc gac acc cgc gaa cgt aaa atc gag	1584		
Ala Phe Asp Glu	Gln Gly Gln Ala Asp Thr Arg Glu Arg Lys Ile Glu			
	515 520 525			
att tgc cgc cgc	gct tac aaa att ttg ctc gaa gag gta ggc ttt ccg	1632		
Ile Cys Arg Arg	Ala Tyr Lys Ile Leu Leu Glu Glu Val Gly Phe Pro			
	530 535 540			
ccg gaa gac atc	atc ttc gac ccg aat atc ttc gcc gtc gcc acc ggt	1680		
Pro Glu Asp Ile	Ile Phe Asp Pro Asn Ile Phe Ala Val Ala Thr Gly			
	545 550 555 560			
att gaa gag cac	aac aac tac gcg cag gac ttt atc ggc gct tgt gaa	1728		
Ile Glu Glu His	Asn Asn Tyr Ala Gln Asp Phe Ile Gly Ala Cys Glu			
	565 570 575			
gac atc aaa cgc	gag ctg ccg cac gcg ctg atc tcc ggc ggc gtg tct	1776		
Asp Ile Lys Arg	Glu Leu Pro His Ala Leu Ile Ser Gly Gly Val Ser			
	580 585 590			
aac gtg tcc ttc	tcg ttt cgc ggc aac gac ccg gta cgt gag gct atc	1824		
Asn Val Ser Phe	Ser Phe Arg Gly Asn Asp Pro Val Arg Glu Ala Ile			
	595 600 605			
cac gcg gta ttc	ctc tac tac gcc atc cgc aac ggc atg gac atg ggc	1872		
His Ala Val Phe	Leu Tyr Tyr Ala Ile Arg Asn Gly Met Asp Met Gly			
	610 615 620			
atc gtc aac gcc	ggg caa ctg gcg att tat gac aac ctg cct gcc gaa	1920		
Ile Val Asn Ala	Gly Gln Leu Ala Ile Tyr Asp Asn Leu Pro Ala Glu			
	625 630 635 640			

100

ctg cgc gat gca gtt gaa gat gtc att ctt aac cgt cgc gat gac ggc Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg Asp Asp Gly 645 650 655	1968
acc gag cgt ttg ctg gat ttg gcg gag aaa tat cgc ggc agc aaa acc Thr Glu Arg Leu Leu Asp Leu Ala Glu Lys Tyr Arg Gly Ser Lys Thr 660 665 670	2016
gac gag gct gcc agt gcc cag cag gcg gaa tgg cgt agc tgg gac gtg Asp Glu Ala Ala Ser Ala Gln Gln Ala Glu Trp Arg Ser Trp Asp Val 675 680 685	2064
aaa aag cgt ctc gaa tac tcg ctg gtg aaa ggc att acc gag ttt atc Lys Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr Glu Phe Ile 690 695 700	2112
gaa cag gat acc gaa gaa gcc cgt cag cag gcc gcc cgc ccg att gag Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Ile Glu 705 710 715 720	2160
gtg att gaa ggg ccg ctg atg gac ggc atg aac gtg gtc ggc gac ctg Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu 725 730 735	2208
ttc ggc gaa ggg aaa atg ttc ctg ccg cag gtg gtg aaa tcc gct cgc Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg 740 745 750	2256
gtg atg aaa caa gcg gtg gcc tac ctg gag ccg ttt att gaa gcc agc Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile Glu Ala Ser 755 760 765	2304
aaa gaa aaa ggc tcc agc aac ggc aag atg gtg att gcc acc gtg aag Lys Glu Lys Gly Ser Ser Asn Gly Lys Met Val Ile Ala Thr Val Lys 770 775 780	2352
ggc gat gtg cac gac att ggc aag aac att gtc ggc gtg gtg ctg caa Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln 785 790 795 800	2400
tgc aac aac tac gaa atc gtc gat ctt ggc gtg atg gtg cca gcg gag Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val Pro Ala Glu 805 810 815	2448
aaa atc ctc aga acg gcg cgt gaa gtg aat gcc gat ctg att ggt ctt Lys Ile Leu Arg Thr Ala Arg Glu Val Asn Ala Asp Leu Ile Gly Leu 820 825 830	2496
tcc ggg ctt atc acc ccg tcg ctg gac gaa atg gtc aac gtg gcg aaa Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn Val Ala Lys 835 840 845	2544
gag atg gag cgt cag gcc ttt act atc ccg cta ctg atc ggc ggc gca Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile Gly Gly Ala 850 855 860	2592
acc act tcg aaa gcg cat acg gcg gtg aaa atc gag cag aac tac agc Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln Asn Tyr Ser 865 870 875 880	2640
ggc ccg acg gtc tac gtg cag aat gct tcg cgt acc gtg ggc gtg gtg Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val Gly Val Val 885 890 895 900	2688

101

885	890	895	
gcg gcg cta ctc tcc gac acc cag cgt gat gac ttt gtc gcc cgt acc			2736
Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val Ala Arg Thr			
900	905	910	
cgc aaa gag tac gaa acc gtg cgt att cag cac gcc cgc aaa aaa ccg			2784
Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Ala Arg Lys Lys Pro			
915	920	925	
cgc acg ccg ccg gtc acg ctg gaa gcg gcg cgc gat aat gat ctg gca			2832
Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn Asp Leu Ala			
930	935	940	
ttt gat tgg gaa cgc tac acc ccg ccg gta gcc cac cgt ctg ggc gtg			2880
Phe Asp Trp Glu Arg Tyr Thr Pro Pro Val Ala His Arg Leu Gly Val			
945	950	955	960
cag gag gtg gaa gcc agc atc gaa acg ctg cgc aac tac atc gac tgg			2928
Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr Ile Asp Trp			
965	970	975	
acg ccg ttc ttt atg acc tgg tcg ctg gcc gcc aaa tac ccg cgc att			2976
Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr Pro Arg Ile			
980	985	990	
ctg gaa gat gag gtg gtg ggc gtt gag gcg cag cgt ctg ttt aaa gac			3024
Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu Phe Lys Asp			
995	1000	1005	
gcc aat gat atg ctg gat aaa ctg agc gcc gag aaa ctg ttg aat ccg			3072
Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Leu Leu Asn Pro			
1010	1015	1020	
cgt ggc gtg gtg ggc ctg ttc ccg gcg aac cgt gtg ggt gac gac atc			3120
Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly Asp Asp Ile			
1025	1030	1035	1040
gaa atc tat cgc gac gaa acc cgt act cat gtt ctg acg gtc agc cac			3168
Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Leu Thr Val Ser His			
1045	1050	1055	
cac ctg cgc cag cag acc gag aaa gtt ggt ttt gct aac tac tgt ctg			3216
His Leu Arg Gln Gln Thr Glu Lys Val Gly Phe Ala Asn Tyr Cys Leu			
1060	1065	1070	
gcg gat ttt gtc gcg ccg aaa ctg agc gcc aaa gcg gac tac atc ggt			3264
Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp Tyr Ile Gly			
1075	1080	1085	
gct ttc gcg gtg acc ggc ggt ctg aag gag gat gcg ctg gcg gac gcc			3312
Ala Phe Ala Val Thr Gly Gly Leu Lys Glu Asp Ala Leu Ala Asp Ala			
1090	1095	1100	
ttc gaa gcg caa cac gac gac tat aac aag atc atg gtg aaa gcg att			3360
Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val Lys Ala Ile			
1105	1110	1115	1120
gcc gac cgt ctg gcg gaa gcg ttt gcc gag tat ctg cat gag cgt gta			3408
Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr Leu His Glu Arg Val			
1125	1130	1135	

102

cgt aag gtt tac tgg gga tat gcg ccg aac gag agc ctg agt aac gac 3456  
 Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Ser Leu Ser Asn Asp  
 1140 1145 1150

gaa tta atc cgc gaa aac tac cag ggg att cgc ccg gcg ccg ggt tat 3504  
 Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr  
 1155 1160 1165

cct gcc tgc ccg gaa cat acc gaa aaa ggc act atc tgg cag cta ctg 3552  
 Pro Ala Cys Pro Glu His Thr Glu Lys Gly Thr Ile Trp Gln Leu Leu  
 1170 1175 1180

gat gtc gaa aaa cac acc ggg atg aag ctc acc gaa tct ttc gcc atg 3600  
 Asp Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser Phe Ala Met  
 1185 1190 1195 1200

tgg cct ggc gcg tgc gtc tcc ggc tgg tac ttc agc cat cct gag agc 3648  
 Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His Pro Glu Ser  
 1205 1210 1215

aaa tac ttc gcg gta gcg cag atc caa cgc gat cag gtg aca gat tat 3696  
 Lys Tyr Phe Ala Val Ala Gln Ile Gln Arg Asp Gln Val Thr Asp Tyr  
 1220 1225 1230

gct ttc cgt aaa gga atg agc gtt gag gac gtt gag cgg tgg ctc gcg 3744  
 Ala Phe Arg Lys Gly Met Ser Val Glu Asp Val Glu Arg Trp Leu Ala  
 1235 1240 1245

ccg aac ctg ggt tac gat gcg gac tga 3771  
 Pro Asn Leu Gly Tyr Asp Ala Asp  
 1250 1255

&lt;210&gt; 24

&lt;211&gt; 1256

&lt;212&gt; PRT

&lt;213&gt; Salmonella typhi

&lt;400&gt; 24

Met Ser His Val Ala Arg Cys Ser Leu Phe Arg Gln His Ala Leu Cys  
 1 5 10 15

Gln Tyr Gly Ser Leu Arg Gly Ala Leu Ser Gly Ala Ser Val Ser Ser  
 20 25 30

Lys Val Glu Gln Leu Arg Ala Gln Leu Asn Glu Arg Ile Leu Val Leu  
 35 40 45

Asp Gly Gly Met Gly Thr Met Ile Gln Ser Tyr Arg Leu His Glu Glu  
 50 55 60

Asp Phe Arg Gly Glu Arg Phe Ala Asp Trp Pro Cys Asp Leu Lys Gly  
 65 70 75 80

Asn Asn Asp Leu Leu Val Leu Ser Lys Pro Glu Val Ile Ala Ala Ile  
 85 90 95

His Asn Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn Thr  
 100 105 110

Phe Asn Ser Thr Thr Ile Ala Met Ala Asp Tyr Arg Met Glu Ser Leu  
 115 120 125

103

Ser Ala Glu Ile Asn Tyr Ala Ala Ala Lys Leu Ala Arg Ala Cys Ala  
 130 135 140  
 Asp Glu Trp Thr Ala Arg Thr Pro Glu Lys Pro Arg Phe Val Ala Gly  
 145 150 155 160  
 Val Leu Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp Val Asn  
 165 170 175  
 Asp Pro Ala Phe Arg Asn Ile Thr Phe Asp Gln Leu Val Ala Ala Tyr  
 180 185 190  
 Arg Glu Ser Thr Lys Ala Leu Val Glu Gly Gly Ala Asp Leu Ile Leu  
 195 200 205  
 Ile Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Val Phe Ala  
 210 215 220  
 Val Lys Glu Glu Phe Glu Ala Leu Gly Val Asp Leu Pro Ile Met Ile  
 225 230 235 240  
 Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr  
 245 250 255  
 Thr Glu Ala Phe Tyr Asn Ser Leu Arg His Ala Glu Ala Leu Thr Phe  
 260 265 270  
 Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu Arg Gln Tyr Val Gln  
 275 280 285  
 Glu Leu Ser Arg Ile Ala Glu Cys Tyr Val Thr Ala His Pro Asn Ala  
 290 295 300  
 Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu Asp Ala Asp Thr Met  
 305 310 315 320  
 Ala Lys Gln Ile Arg Glu Trp Ala Glu Ala Gly Phe Leu Asn Ile Val  
 325 330 335  
 Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Ala Ala Met Ser Arg  
 340 345 350  
 Ala Val Ala Gly Leu Ser Pro Arg Gln Leu Pro Asp Ile Pro Val Ala  
 355 360 365  
 Cys Arg Leu Ser Gly Leu Glu Pro Leu Asn Ile Gly Asp Asp Ser Leu  
 370 375 380  
 Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Lys Phe  
 385 390 395 400  
 Lys Arg Leu Ile Lys Glu Glu Lys Tyr Ser Glu Ala Leu Asp Val Ala  
 405 410 415  
 Arg Gln Gln Val Glu Ser Gly Ala Gln Ile Ile Asp Ile Asn Met Asp  
 420 425 430  
 Glu Gly Met Leu Asp Ala Glu Ala Ala Met Val Arg Phe Leu Ser Leu  
 435 440 445  
 Ile Ala Gly Glu Pro Asp Ile Ala Arg Val Pro Ile Met Ile Asp Ser

104

450                      455                      460  
 Ser Lys Trp Glu Val Ile Glu Lys Gly Leu Lys Cys Ile Gln Gly Lys  
 465                      470                      475                      480  
 Gly Ile Val Asn Ser Ile Ser Met Lys Glu Gly Val Glu Ala Phe Ile  
                     485                      490                      495  
 His His Ala Lys Leu Leu Arg Arg Tyr Gly Ala Ala Val Val Val Met  
                     500                      505                      510  
 Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Arg Glu Arg Lys Ile Glu  
                     515                      520                      525  
 Ile Cys Arg Arg Ala Tyr Lys Ile Leu Leu Glu Glu Val Gly Phe Pro  
                     530                      535                      540  
 Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe Ala Val Ala Thr Gly  
 545                      550                      555                      560  
 Ile Glu Glu His Asn Asn Tyr Ala Gln Asp Phe Ile Gly Ala Cys Glu  
                     565                      570                      575  
 Asp Ile Lys Arg Glu Leu Pro His Ala Leu Ile Ser Gly Gly Val Ser  
                     580                      585                      590  
 Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Pro Val Arg Glu Ala Ile  
                     595                      600                      605  
 His Ala Val Phe Leu Tyr Tyr Ala Ile Arg Asn Gly Met Asp Met Gly  
                     610                      615                      620  
 Ile Val Asn Ala Gly Gln Leu Ala Ile Tyr Asp Asn Leu Pro Ala Glu  
 625                      630                      635                      640  
 Leu Arg Asp Ala Val Glu Asp Val Ile Leu Asn Arg Arg Asp Asp Gly  
                     645                      650                      655  
 Thr Glu Arg Leu Leu Asp Leu Ala Glu Lys Tyr Arg Gly Ser Lys Thr  
                     660                      665                      670  
 Asp Glu Ala Ala Ser Ala Gln Gln Ala Glu Trp Arg Ser Trp Asp Val  
                     675                      680                      685  
 Lys Lys Arg Leu Glu Tyr Ser Leu Val Lys Gly Ile Thr Glu Phe Ile  
                     690                      695                      700  
 Glu Gln Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Ile Glu  
 705                      710                      715                      720  
 Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu  
                     725                      730                      735  
 Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg  
                     740                      745                      750  
 Val Met Lys Gln Ala Val Ala Tyr Leu Glu Pro Phe Ile Glu Ala Ser  
                     755                      760                      765  
 Lys Glu Lys Gly Ser Ser Asn Gly Lys Met Val Ile Ala Thr Val Lys  
                     770                      775                      780

105

Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln  
 785 790 795 800  
 Cys Asn Asn Tyr Glu Ile Val Asp Leu Gly Val Met Val Pro Ala Glu  
 805 810 815  
 Lys Ile Leu Arg Thr Ala Arg Glu Val Asn Ala Asp Leu Ile Gly Leu  
 820 825 830  
 Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Asn Val Ala Lys  
 835 840 845  
 Glu Met Glu Arg Gln Gly Phe Thr Ile Pro Leu Leu Ile Gly Gly Ala  
 850 855 860  
 Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile Glu Gln Asn Tyr Ser  
 865 870 875 880  
 Gly Pro Thr Val Tyr Val Gln Asn Ala Ser Arg Thr Val Gly Val Val  
 885 890 895  
 Ala Ala Leu Leu Ser Asp Thr Gln Arg Asp Asp Phe Val Ala Arg Thr  
 900 905 910  
 Arg Lys Glu Tyr Glu Thr Val Arg Ile Gln His Ala Arg Lys Lys Pro  
 915 920 925  
 Arg Thr Pro Pro Val Thr Leu Glu Ala Ala Arg Asp Asn Asp Leu Ala  
 930 935 940  
 Phe Asp Trp Glu Arg Tyr Thr Pro Pro Val Ala His Arg Leu Gly Val  
 945 950 955 960  
 Gln Glu Val Glu Ala Ser Ile Glu Thr Leu Arg Asn Tyr Ile Asp Trp  
 965 970 975  
 Thr Pro Phe Phe Met Thr Trp Ser Leu Ala Gly Lys Tyr Pro Arg Ile  
 980 985 990  
 Leu Glu Asp Glu Val Val Gly Val Glu Ala Gln Arg Leu Phe Lys Asp  
 995 1000 1005  
 Ala Asn Asp Met Leu Asp Lys Leu Ser Ala Glu Lys Leu Leu Asn Pro  
 1010 1015 1020  
 Arg Gly Val Val Gly Leu Phe Pro Ala Asn Arg Val Gly Asp Asp Ile  
 1025 1030 1035 1040  
 Glu Ile Tyr Arg Asp Glu Thr Arg Thr His Val Leu Thr Val Ser His  
 1045 1050 1055  
 His Leu Arg Gln Gln Thr Glu Lys Val Gly Phe Ala Asn Tyr Cys Leu  
 1060 1065 1070  
 Ala Asp Phe Val Ala Pro Lys Leu Ser Gly Lys Ala Asp Tyr Ile Gly  
 1075 1080 1085  
 Ala Phe Ala Val Thr Gly Gly Leu Lys Glu Asp Ala Leu Ala Asp Ala  
 1090 1095 1100  
 Phe Glu Ala Gln His Asp Asp Tyr Asn Lys Ile Met Val Lys Ala Ile  
 1105 1110 1115 1120

Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Tyr Leu His Glu Arg Val  
 1125 1130 1135

Arg Lys Val Tyr Trp Gly Tyr Ala Pro Asn Glu Ser Leu Ser Asn Asp  
 1140 1145 1150

Glu Leu Ile Arg Glu Asn Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr  
 1155 1160 1165

Pro Ala Cys Pro Glu His Thr Glu Lys Gly Thr Ile Trp Gln Leu Leu  
 1170 1175 1180

Asp Val Glu Lys His Thr Gly Met Lys Leu Thr Glu Ser Phe Ala Met  
 1185 1190 1195 1200

Trp Pro Gly Ala Ser Val Ser Gly Trp Tyr Phe Ser His Pro Glu Ser  
 1205 1210 1215

Lys Tyr Phe Ala Val Ala Gln Ile Gln Arg Asp Gln Val Thr Asp Tyr  
 1220 1225 1230

Ala Phe Arg Lys Gly Met Ser Val Glu Asp Val Glu Arg Trp Leu Ala  
 1235 1240 1245

Pro Asn Leu Gly Tyr Asp Ala Asp  
 1250 1255

<210> 25  
 <211> 3711  
 <212> DNA  
 <213> Pseudomonas fluorescens

<220>  
 <221> CDS  
 <222> (1) .. (3708)  
 <223> RPU03563

<400> 25  
 atg tcc gat cgc agc gtc cgc ctt caa gct ctc aag caa gct ctc aaa 48  
 Met Ser Asp Arg Ser Val Arg Leu Gln Ala Leu Lys Gln Ala Leu Lys  
 1 5 10 15  
 gag cgc atc ctg att ctc gac ggc ggc atg ggc acg atg atc cag agc 96  
 Glu Arg Ile Leu Ile Leu Asp Gly Gly Met Gly Thr Met Ile Gln Ser  
 20 25 30  
 tac aag ctc gaa gag cag gat tat cgc ggc aaa cgc ttc gcc gac tgg 144  
 Tyr Lys Leu Glu Glu Gln Asp Tyr Arg Gly Lys Arg Phe Ala Asp Trp  
 35 40 45  
 ccg agc gac gtc aag ggc aac aac gac ctg ttg gtg ctg acc cgc ccg 192  
 Pro Ser Asp Val Lys Gly Asn Asn Asp Leu Leu Val Leu Thr Arg Pro  
 50 55 60  
 gac gtg atc ggc ggc atc gag aaa gcc tat ctg gat gcc ggt gcc gac 240  
 Asp Val Ile Gly Gly Ile Glu Lys Ala Tyr Leu Asp Ala Gly Ala Asp  
 65 70 75 80  
 atc ctc gag acc aac acc ttc aac gcc acg cag att tcc atg gcc gac 288  
 Ile Leu Glu Thr Asn Thr Phe Asn Ala Thr Gln Ile Ser Met Ala Asp



										107				
85					90					95				
tac ggc atg gaa gaa ctg gtc tac gaa ctc aac gta gaa ggc gcc cgt	336													
Tyr Gly Met Glu Glu Leu Val Tyr Glu Leu Asn Val Glu Gly Ala Arg														
100	105	110												
ctg gca cgc aag gtc gcc gac gcg aaa acc ctc gag acc ccc gac aag	384													
Leu Ala Arg Lys Val Ala Asp Ala Lys Thr Leu Glu Thr Pro Asp Lys														
115	120	125												
ccg cgc ttc gtc gcc ggc gtt ctc ggc ccg acc agc cgc acc tgc tcg	432													
Pro Arg Phe Val Ala Gly Val Leu Gly Pro Thr Ser Arg Thr Cys Ser														
130	135	140												
ctg tcg ccg gac gtc aac aac ccg ggc tat cgc aac gtc acc ttc gat	480													
Leu Ser Pro Asp Val Asn Asn Pro Gly Tyr Arg Asn Val Thr Phe Asp														
145	150	155	160											
gag ctg gtc gaa aac tac acc gag gcc acc aaa ggc ctg atc gag ggc	528													
Glu Leu Val Glu Asn Tyr Thr Glu Ala Thr Lys Gly Leu Ile Glu Gly														
165	170	175												
ggc gcg gat ctg atc ctg atc gaa acc atc ttc gac acc ctc aac gcc	576													
Gly Ala Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala														
180	185	190												
aaa gcc gcg atc ttc gcc gtg caa ggc gtg ttc gaa gaa ctg ggc ttc	624													
Lys Ala Ala Ile Phe Ala Val Gln Gly Val Phe Glu Glu Leu Gly Phe														
195	200	205												
gaa ttg ccg atc atg atc tcc ggc acc atc acc gac gcc tcc ggc cgt	672													
Glu Leu Pro Ile Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg														
210	215	220												
acc ctg tcg ggc cag acc acc gaa gcg ttc tgg aac tcc gtg gct cac	720													
Thr Leu Ser Gly Gln Thr Thr Glu Ala Phe Trp Asn Ser Val Ala His														
225	230	235	240											
gcc aaa ccg att tcc gtc ggt ctt aac tgc gcc ctc ggc gcc cgc gaa	768													
Ala Lys Pro Ile Ser Val Gly Leu Asn Cys Ala Leu Gly Ala Arg Glu														
245	250	255												
ctg cgt ccg tac ctg gaa gag ctg tcg gac aag gcc agc acc cac gtt	816													
Leu Arg Pro Tyr Leu Glu Glu Leu Ser Asp Lys Ala Ser Thr His Val														
260	265	270												
tcg gcg cac ccg aac gcc ggc ctg ccg aac gaa ttc ggc gag tac gac	864													
Ser Ala His Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp														
275	280	285												
gag ctg ccg gtg gac acc gcc aag gtc atc gaa gag ttc gcc cag agc	912													
Glu Leu Pro Val Asp Thr Ala Lys Val Ile Glu Glu Phe Ala Gln Ser														
290	295	300												
ggt ttc ctc aac atc gtc ggc ggt tgc tgc ggc acc acg ccg ggc cat	960													
Gly Phe Leu Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Gly His														
305	310	315	320											
atc gaa gcc atc gcc aaa gcc gtt gcc ggt tac gcg cca cgg cag att	1008													
Ile Glu Ala Ile Ala Lys Ala Val Ala Gly Tyr Ala Pro Arg Gln Ile														
325	330	335												

108

ccg gac att ccc aag gcc tgc cgc ctg tcg ggt ctg gaa ccg ttc acc	1056
Pro Asp Ile Pro Lys Ala Cys Arg Leu Ser Gly Leu Glu Pro Phe Thr	
340 345 350	
att gat cgc agc tcg ctg ttc gtc aac gtc ggc gag cgg acc aac atc	1104
Ile Asp Arg Ser Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn Ile	
355 360 365	
acc ggg tcc gcg aaa ttt gcc cgg ctg atc cgt gaa gac aac tac acc	1152
Thr Gly Ser Ala Lys Phe Ala Arg Leu Ile Arg Glu Asp Asn Tyr Thr	
370 375 380	
gaa gcc ctg gaa gtc gcc ctg cag cag gtc gag gcc ggc gcc cag gtg	1200
Glu Ala Leu Glu Val Ala Leu Gln Gln Val Glu Ala Gly Ala Gln Val	
385 390 395 400	
atc gac atc aac atg gac gaa ggg atg ctc gat tcg aag aag gcc atg	1248
Ile Asp Ile Asn Met Asp Glu Gly Met Leu Asp Ser Lys Lys Ala Met	
405 410 415	
gtg acc ttc ctc aat ctg att gcc ggc gaa ccg gac atc tcc cgc gta	1296
Val Thr Phe Leu Asn Leu Ile Ala Gly Glu Pro Asp Ile Ser Arg Val	
420 425 430	
ccg atc atg atc gac tcc tcg aaa tgg gac gtg atc gaa gcc ggc ctc	1344
Pro Ile Met Ile Asp Ser Ser Lys Trp Asp Val Ile Glu Ala Gly Leu	
435 440 445	
aag tgc att cag ggc aag ggc atc gtc aac tcg atc agc atg aaa gaa	1392
Lys Cys Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Met Lys Glu	
450 455 460	
ggc gtc gag cag ttc atc cac cac gcc aaa ctg tgc aag cgc tat ggc	1440
Gly Val Glu Gln Phe Ile His His Ala Lys Leu Cys Lys Arg Tyr Gly	
465 470 475 480	
gcc gcc gtg gtg gtg atg gcg ttc gac gaa gcc ggc cag gct gac acc	1488
Ala Ala Val Val Val Met Ala Phe Asp Glu Ala Gly Gln Ala Asp Thr	
485 490 495	
gaa gcg cgc aag aaa gag atc tgc aaa cgc tcc tac gac att ctg gtc	1536
Glu Ala Arg Lys Lys Glu Ile Cys Lys Arg Ser Tyr Asp Ile Leu Val	
500 505 510	
aac gaa gtc ggc ttc ccg ccg gaa gac atc att ttc gac ccg aac atc	1584
Asn Glu Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile	
515 520 525	
ttc gcc gtg gcc acc ggc atc gaa gaa cac aac aac tac gct gtg gac	1632
Phe Ala Val Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Val Asp	
530 535 540	
ttc atc aac gcc tgt gcc tac atc cgc gac gag ctg ccg tat gcc ctg	1680
Phe Ile Asn Ala Cys Ala Tyr Ile Arg Asp Glu Leu Pro Tyr Ala Leu	
545 550 555 560	
agc tcc ggc ggc gtg tcc aac gtg tcg ttc tcg ttc cgc ggc aac aac	1728
Ser Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn	
565 570 575	
ccg gtg cgc gag gcg atc cac tcg gtg ttc ctg ctg tac gcg atc cgc	1776
Pro Val Arg Glu Ala Ile His Ser Val Phe Leu Leu Tyr Ala Ile Arg	

109

580	585	590	
gcc ggc ctg acc atg ggt atc gtc aac gcc ggt cag ctg gag atc tac Ala Gly Leu Thr Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr 595 600 605			1824
gac cag atc ccg cag gaa ctg cgc gac gcc gtt gaa gac gtg atc ctc Asp Gln Ile Pro Gln Glu Leu Arg Asp Ala Val Glu Asp Val Ile Leu 610 615 620			1872
aac cgc acg ccg gaa ggc acc gac gcc ctc ctc gcc atc gcc gac aag Asn Arg Thr Pro Glu Gly Thr Asp Ala Leu Leu Ala Ile Ala Asp Lys 625 630 635 640			1920
tac aag ggc gac ggc agc gtc aag gaa gcc gag acc gaa gaa tgg cgc Tyr Lys Gly Asp Gly Ser Val Lys Glu Ala Glu Thr Glu Glu Trp Arg 645 650 655			1968
ggc tgg gac gtc aac aaa cgt ctg gaa cat gcg ctg gtc aag ggc atc Gly Trp Asp Val Asn Lys Arg Leu Glu His Ala Leu Val Lys Gly Ile 660 665 670			2016
acc acc cac atc gtc gaa gac acc gaa gaa tcc cgt cag tcc ttc gcc Thr Thr His Ile Val Glu Asp Thr Glu Glu Ser Arg Gln Ser Phe Ala 675 680 685			2064
cgc ccg atc gaa gtg atc gaa ggc ccg ctg atg tcc ggc atg aac atc Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Ser Gly Met Asn Ile 690 695 700			2112
gtc ggc gac ctg ttc ggc gcc ggc aaa atg ttc ctg ccg caa gtg gtg Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln Val Val 705 710 715 720			2160
aaa tcc gcc cgc gtg atg aag cag gcc gtg gcg cac ctg att ccg ttc Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Ile Pro Phe 725 730 735			2208
atc gaa ctg gaa aaa ggc gac aag ccg gaa gcc aag ggc aag atc ctg Ile Glu Leu Glu Lys Gly Asp Lys Pro Glu Ala Lys Gly Lys Ile Leu 740 745 750			2256
atg gcc acg gtc aaa ggc gac gtg cac gac atc ggc aag aac atc gtc Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val 755 760 765			2304
ggc gtg gtg ctg ggt tgc aac ggc tac gac atc gtc gac ctc ggc gtg Gly Val Val Leu Gly Cys Asn Gly Tyr Asp Ile Val Asp Leu Gly Val 770 775 780			2352
atg gtg ccg gcg gag aag atc ctg cag gtg gcc aag gag cag aag tgc Met Val Pro Ala Glu Lys Ile Leu Gln Val Ala Lys Glu Gln Lys Cys 785 790 795 800			2400
gac atc atc ggc ctg tcc ggt ctg atc acc ccg tcg ctg gat gag atg Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met 805 810 815			2448
gtc cat gtg gcc cgc gag atg cag cgc cag gac ttc cac ctg ccg ctg Val His Val Ala Arg Glu Met Gln Arg Gln Asp Phe His Leu Pro Leu 820 825 830			2496

110

atg atc ggc ggc gcg acc acc tcc aag gcg cac acg gcg gtg aag atc Met Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile 835 840 845	2544
gag ccc aag tac agc aac gac gca gtg gtc tac gtg acc gac gcc tcc Glu Pro Lys Tyr Ser Asn Asp Ala Val Val Tyr Val Thr Asp Ala Ser 850 855 860	2592
cgc gcc gtg ggc gtg gcg acg cag ttg ctg tcc aag gaa ctg aaa gcc Arg Ala Val Gly Val Ala Thr Gln Leu Leu Ser Lys Glu Leu Lys Ala 865 870 875 880	2640
ggt ttc gtc cag aag acc cgc gaa gag tac atc gac gtc cgc gag cgc Gly Phe Val Gln Lys Thr Arg Glu Glu Tyr Ile Asp Val Arg Glu Arg 885 890 895	2688
acc gcc aac cgc agc gcc cgc acc gaa cgc ctg agc tac gcc gcc gcg Thr Ala Asn Arg Ser Ala Arg Thr Glu Arg Leu Ser Tyr Ala Ala Ala 900 905 910	2736
atc gcc aag aag ccg cag ttc gac tgg gcc act tac acc ccg gtc aaa Ile Ala Lys Lys Pro Gln Phe Asp Trp Ala Thr Tyr Thr Pro Val Lys 915 920 925	2784
ccg acc ttc acc ggc acc cgc gtg ctg gac aac atc gac ctc aac gtt Pro Thr Phe Thr Gly Thr Arg Val Leu Asp Asn Ile Asp Leu Asn Val 930 935 940	2832
ctc gcc gag tac atc gac tgg acg ccg ttc ttc atc tcc tgg gac ctg Leu Ala Glu Tyr Ile Asp Trp Thr Pro Phe Phe Ile Ser Trp Asp Leu 945 950 955 960	2880
gcc ggc aag ttc ccg cgc atc ctc gaa gac gaa gtg gtc ggc gaa gcg Ala Gly Lys Phe Pro Arg Ile Leu Glu Asp Glu Val Val Gly Glu Ala 965 970 975	2928
gcg acc gcg ctg tac aag gac gct cgc gag atg ctg acc aag ctg atc Ala Thr Ala Leu Tyr Lys Asp Ala Arg Glu Met Leu Thr Lys Leu Ile 980 985 990	2976
gac gag aaa ctg atc agc gcc cgt gcg gtg ttc ggc ttc tgg ccg gcc Asp Glu Lys Leu Ile Ser Ala Arg Ala Val Phe Gly Phe Trp Pro Ala 995 1000 1005	3024
aat cag gtg cac gac gac gat atc gag ctg tac ggc gat gac ggc aag Asn Gln Val His Asp Asp Asp Ile Glu Leu Tyr Gly Asp Asp Gly Lys 1010 1015 1020	3072
cca atg gcg cgc ctg cat cac ctg cgc cag cag atc atc aag acc gac Pro Met Ala Arg Leu His His Leu Arg Gln Gln Ile Ile Lys Thr Asp 1025 1030 1035 1040	3120
ggc aaa ccg aac ttc tcc ctc gcc gac ttc gtc gcg ccg aag gac agc Gly Lys Pro Asn Phe Ser Leu Ala Asp Phe Val Ala Pro Lys Asp Ser 1045 1050 1055	3168
gaa gtg acc gac tac gtt ggt ggt ttc atc acc acc gcc ggg atc ggc Glu Val Thr Asp Tyr Val Gly Gly Phe Ile Thr Thr Ala Gly Ile Gly 1060 1065 1070	3216
gcc gaa gaa gtg gcc aag gcc tat cag gac gcc ggc gac gat tac aac Ala Glu Glu Val Ala Lys Ala Tyr Gln Asp Ala Gly Asp Asp Tyr Asn	3264

111

1075

1080

1085

tcg atc atg gtc aag gcc ctg gcc gac cgt ctg gcc gag gcg tgc gcc 3312  
 Ser Ile Met Val Lys Ala Leu Ala Asp Arg Leu Ala Glu Ala Cys Ala  
 1090 1095 1100

gag tgg ctg cac cag cag gtg cgc aaa gag cac tgg ggt tac gcc aag 3360  
 Glu Trp Leu His Gln Gln Val Arg Lys Glu His Trp Gly Tyr Ala Lys  
 1105 1110 1115 1120

gat gaa gcc ctc gat aac gag gcg ctg atc aaa gag cag tat tcc gcc 3408  
 Asp Glu Ala Leu Asp Asn Glu Ala Leu Ile Lys Glu Gln Tyr Ser Gly  
 1125 1130 1135

atc cgc cct gcc ccc gcc tac ccg gcg tgc ccg gat cac acc gag aag 3456  
 Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys  
 1140 1145 1150

gcc acc ctg ttc gcc ctg ctc gac cct gaa gca cag gaa atg cgc gcc 3504  
 Ala Thr Leu Phe Ala Leu Leu Asp Pro Glu Ala Gln Glu Met Arg Ala  
 1155 1160 1165

ggc cgc agc ggt gtg ttc ctc acc gag cac tac gcg atg ttc ccg gcg 3552  
 Gly Arg Ser Gly Val Phe Leu Thr Glu His Tyr Ala Met Phe Pro Ala  
 1170 1175 1180

gca gcc gtc agc gcc tgg tac ttc gcc cat ccg cag gcg cag tac ttc 3600  
 Ala Ala Val Ser Gly Trp Tyr Phe Ala His Pro Gln Ala Gln Tyr Phe  
 1185 1190 1195 1200

gcc gtg gcc aag gtc gac aag gat cag gtg cag agc tac acc tcg cgc 3648  
 Ala Val Gly Lys Val Asp Lys Asp Gln Val Gln Ser Tyr Thr Ser Arg  
 1205 1210 1215

aaa ggc cag gaa ctg agc ctg acc gag cgc tgg ctg gca ccc aat ctg 3696  
 Lys Gly Gln Glu Leu Ser Leu Thr Glu Arg Trp Leu Ala Pro Asn Leu  
 1220 1225 1230

ggc tac gac aac tga 3711  
 Gly Tyr Asp Asn  
 1235

&lt;210&gt; 26

&lt;211&gt; 1236

&lt;212&gt; PRT

<213> *Pseudomonas fluorescens*

&lt;400&gt; 26

Met Ser Asp Arg Ser Val Arg Leu Gln Ala Leu Lys Gln Ala Leu Lys  
 1 5 10 15

Glu Arg Ile Leu Ile Leu Asp Gly Gly Met Gly Thr Met Ile Gln Ser  
 20 25 30

Tyr Lys Leu Glu Glu Gln Asp Tyr Arg Gly Lys Arg Phe Ala Asp Trp  
 35 40 45

Pro Ser Asp Val Lys Gly Asn Asn Asp Leu Leu Val Leu Thr Arg Pro  
 50 55 60

Asp Val Ile Gly Gly Ile Glu Lys Ala Tyr Leu Asp Ala Gly Ala Asp

112

65	70	75	80
Ile Leu Glu Thr Asn Thr Phe Asn Ala Thr Gln Ile Ser Met Ala Asp	85	90	95
Tyr Gly Met Glu Glu Leu Val Tyr Glu Leu Asn Val Glu Gly Ala Arg	100	105	110
Leu Ala Arg Lys Val Ala Asp Ala Lys Thr Leu Glu Thr Pro Asp Lys	115	120	125
Pro Arg Phe Val Ala Gly Val Leu Gly Pro Thr Ser Arg Thr Cys Ser	130	135	140
Leu Ser Pro Asp Val Asn Asn Pro Gly Tyr Arg Asn Val Thr Phe Asp	145	150	155
Glu Leu Val Glu Asn Tyr Thr Glu Ala Thr Lys Gly Leu Ile Glu Gly	165	170	175
Gly Ala Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala	180	185	190
Lys Ala Ala Ile Phe Ala Val Gln Gly Val Phe Glu Glu Leu Gly Phe	195	200	205
Glu Leu Pro Ile Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg	210	215	220
Thr Leu Ser Gly Gln Thr Thr Glu Ala Phe Trp Asn Ser Val Ala His	225	230	235
Ala Lys Pro Ile Ser Val Gly Leu Asn Cys Ala Leu Gly Ala Arg Glu	245	250	255
Leu Arg Pro Tyr Leu Glu Glu Leu Ser Asp Lys Ala Ser Thr His Val	260	265	270
Ser Ala His Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp	275	280	285
Glu Leu Pro Val Asp Thr Ala Lys Val Ile Glu Glu Phe Ala Gln Ser	290	295	300
Gly Phe Leu Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Gly His	305	310	315
Ile Glu Ala Ile Ala Lys Ala Val Ala Gly Tyr Ala Pro Arg Gln Ile	325	330	335
Pro Asp Ile Pro Lys Ala Cys Arg Leu Ser Gly Leu Glu Pro Phe Thr	340	345	350
Ile Asp Arg Ser Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn Ile	355	360	365
Thr Gly Ser Ala Lys Phe Ala Arg Leu Ile Arg Glu Asp Asn Tyr Thr	370	375	380
Glu Ala Leu Glu Val Ala Leu Gln Gln Val Glu Ala Gly Ala Gln Val	385	390	395
			400

113

Ile Asp Ile Asn Met Asp Glu Gly Met Leu Asp Ser Lys Lys Ala Met  
 405 410 415  
 Val Thr Phe Leu Asn Leu Ile Ala Gly Glu Pro Asp Ile Ser Arg Val  
 420 425 430  
 Pro Ile Met Ile Asp Ser Ser Lys Trp Asp Val Ile Glu Ala Gly Leu  
 435 440 445  
 Lys Cys Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Met Lys Glu  
 450 455 460  
 Gly Val Glu Gln Phe Ile His His Ala Lys Leu Cys Lys Arg Tyr Gly  
 465 470 475 480  
 Ala Ala Val Val Val Met Ala Phe Asp Glu Ala Gly Gln Ala Asp Thr  
 485 490 495  
 Glu Ala Arg Lys Lys Glu Ile Cys Lys Arg Ser Tyr Asp Ile Leu Val  
 500 505 510  
 Asn Glu Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile  
 515 520 525  
 Phe Ala Val Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Val Asp  
 530 535 540  
 Phe Ile Asn Ala Cys Ala Tyr Ile Arg Asp Glu Leu Pro Tyr Ala Leu  
 545 550 555 560  
 Ser Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn  
 565 570 575  
 Pro Val Arg Glu Ala Ile His Ser Val Phe Leu Leu Tyr Ala Ile Arg  
 580 585 590  
 Ala Gly Leu Thr Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr  
 595 600 605  
 Asp Gln Ile Pro Gln Glu Leu Arg Asp Ala Val Glu Asp Val Ile Leu  
 610 615 620  
 Asn Arg Thr Pro Glu Gly Thr Asp Ala Leu Leu Ala Ile Ala Asp Lys  
 625 630 635 640  
 Tyr Lys Gly Asp Gly Ser Val Lys Glu Ala Glu Thr Glu Glu Trp Arg  
 645 650 655  
 Gly Trp Asp Val Asn Lys Arg Leu Glu His Ala Leu Val Lys Gly Ile  
 660 665 670  
 Thr Thr His Ile Val Glu Asp Thr Glu Glu Ser Arg Gln Ser Phe Ala  
 675 680 685  
 Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Ser Gly Met Asn Ile  
 690 695 700  
 Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln Val Val  
 705 710 715 720  
 Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Ile Pro Phe  
 725 730 735

114

Ile Glu Leu Glu Lys Gly Asp Lys Pro Glu Ala Lys Gly Lys Ile Leu  
 740 745 750  
 Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val  
 755 760 765  
 Gly Val Val Leu Gly Cys Asn Gly Tyr Asp Ile Val Asp Leu Gly Val  
 770 775 780  
 Met Val Pro Ala Glu Lys Ile Leu Gln Val Ala Lys Glu Gln Lys Cys  
 785 790 795 800  
 Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met  
 805 810 815  
 Val His Val Ala Arg Glu Met Gln Arg Gln Asp Phe His Leu Pro Leu  
 820 825 830  
 Met Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile  
 835 840 845  
 Glu Pro Lys Tyr Ser Asn Asp Ala Val Val Tyr Val Thr Asp Ala Ser  
 850 855 860  
 Arg Ala Val Gly Val Ala Thr Gln Leu Leu Ser Lys Glu Leu Lys Ala  
 865 870 875 880  
 Gly Phe Val Gln Lys Thr Arg Glu Glu Tyr Ile Asp Val Arg Glu Arg  
 885 890 895  
 Thr Ala Asn Arg Ser Ala Arg Thr Glu Arg Leu Ser Tyr Ala Ala Ala  
 900 905 910  
 Ile Ala Lys Lys Pro Gln Phe Asp Trp Ala Thr Tyr Thr Pro Val Lys  
 915 920 925  
 Pro Thr Phe Thr Gly Thr Arg Val Leu Asp Asn Ile Asp Leu Asn Val  
 930 935 940  
 Leu Ala Glu Tyr Ile Asp Trp Thr Pro Phe Phe Ile Ser Trp Asp Leu  
 945 950 955 960  
 Ala Gly Lys Phe Pro Arg Ile Leu Glu Asp Glu Val Val Gly Glu Ala  
 965 970 975  
 Ala Thr Ala Leu Tyr Lys Asp Ala Arg Glu Met Leu Thr Lys Leu Ile  
 980 985 990  
 Asp Glu Lys Leu Ile Ser Ala Arg Ala Val Phe Gly Phe Trp Pro Ala  
 995 1000 1005  
 Asn Gln Val His Asp Asp Asp Ile Glu Leu Tyr Gly Asp Asp Gly Lys  
 1010 1015 1020  
 Pro Met Ala Arg Leu His His Leu Arg Gln Gln Ile Ile Lys Thr Asp  
 1025 1030 1035 1040  
 Gly Lys Pro Asn Phe Ser Leu Ala Asp Phe Val Ala Pro Lys Asp Ser  
 1045 1050 1055  
 Glu Val Thr Asp Tyr Val Gly Gly Phe Ile Thr Thr Ala Gly Ile Gly



115

1060 1065 1070  
 Ala Glu Glu Val Ala Lys Ala Tyr Gln Asp Ala Gly Asp Asp Tyr Asn  
 1075 1080 1085  
 Ser Ile Met Val Lys Ala Leu Ala Asp Arg Leu Ala Glu Ala Cys Ala  
 1090 1095 1100  
 Glu Trp Leu His Gln Gln Val Arg Lys Glu His Trp Gly Tyr Ala Lys  
 1105 1110 1115 1120  
 Asp Glu Ala Leu Asp Asn Glu Ala Leu Ile Lys Glu Gln Tyr Ser Gly  
 1125 1130 1135  
 Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys  
 1140 1145 1150  
 Ala Thr Leu Phe Ala Leu Leu Asp Pro Glu Ala Gln Glu Met Arg Ala  
 1155 1160 1165  
 Gly Arg Ser Gly Val Phe Leu Thr Glu His Tyr Ala Met Phe Pro Ala  
 1170 1175 1180  
 Ala Ala Val Ser Gly Trp Tyr Phe Ala His Pro Gln Ala Gln Tyr Phe  
 1185 1190 1195 1200  
 Ala Val Gly Lys Val Asp Lys Asp Gln Val Gln Ser Tyr Thr Ser Arg  
 1205 1210 1215  
 Lys Gly Gln Glu Leu Ser Leu Thr Glu Arg Trp Leu Ala Pro Asn Leu  
 1220 1225 1230  
 Gly Tyr Asp Asn  
 1235  
  
 <210> 27  
 <211> 3705  
 <212> DNA  
 <213> *Pseudomonas aeruginosa*  
  
 <220>  
 <221> CDS  
 <222> (1) .. (3702)  
 <223> RPA01772  
  
 <400> 27  
 atg tcc agc ccg ctc acc gat cgc agc gcc cgc ctg caa gcc ctc cag 48  
 Met Ser Ser Pro Leu Thr Asp Arg Ser Ala Arg Leu Gln Ala Leu Gln  
 1 5 10 15  
 cac gcc ctc agg gaa cgt atc ctg atc ctc gat ggc ggc atg ggc acc 96  
 His Ala Leu Arg Glu Arg Ile Leu Ile Leu Asp Gly Gly Met Gly Thr  
 20 25 30  
 atg atc cag agc tac aag ctg gaa gag gcc gac tac cgc ggc gag cgc 144  
 Met Ile Gln Ser Tyr Lys Leu Glu Glu Ala Asp Tyr Arg Gly Glu Arg  
 35 40 45  
 ttc gcc gac tgg ccg agc gac gtg aaa ggc aac aac gac ctc ttg ctg 192  
 Phe Ala Asp Trp Pro Ser Asp Val Lys Gly Asn Asn Asp Leu Leu Leu  
 50 55 60

116

ctg agc cgc ccg gac gtg atc cag gcc atc gag aag gcc tac ctc gac Leu Ser Arg Pro Asp Val Ile Gln Ala Ile Glu Lys Ala Tyr Leu Asp 65 70 75 80	240
gcc ggc gcc gac atc ctc gag acc aac acc ttc aac gcc acc cag gtg Ala Gly Ala Asp Ile Leu Glu Thr Asn Thr Phe Asn Ala Thr Gln Val 85 90 95	288
tcc cag gcc gac tac ggc atg cag tgc ctg gcc tac gaa ctc aac gtc Ser Gln Ala Asp Tyr Gly Met Gln Ser Leu Ala Tyr Glu Leu Asn Val 100 105 110	336
gaa ggc gcc cgc ctg gcc cgc cag gtg gcg gac gcg aag acc gcc gag Glu Gly Ala Arg Leu Ala Arg Gln Val Ala Asp Ala Lys Thr Ala Glu 115 120 125	384
acc ccg gac aag ccg cgt ttc gtc gcc ggc gtg ctc ggc ccg acc agc Thr Pro Asp Lys Pro Arg Phe Val Ala Gly Val Leu Gly Pro Thr Ser 130 135 140	432
cgc acc tgc tgc att tcc ccg gac gtg aac aac ccc ggc tac cgc aac Arg Thr Cys Ser Ile Ser Pro Asp Val Asn Asn Pro Gly Tyr Arg Asn 145 150 155 160	480
gtc acc ttc gac gaa ctg gtg gag aac tac gtc gag gcg acc cga ggc Val Thr Phe Asp Glu Leu Val Glu Asn Tyr Val Glu Ala Thr Arg Gly 165 170 175	528
ctg atc gaa ggc ggc gcc gac ctg atc ctg atc gag acc atc ttc gac Leu Ile Glu Gly Ala Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp 180 185 190	576
acc ctc aac gcc aag gcg gcg atc ttc gcc gtc cag ggc gtg ttc gag Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Val Gln Gly Val Phe Glu 195 200 205	624
gaa ctc ggc gtg gag ctg ccg atc atg atc tcc gga acc atc acc gac Glu Leu Gly Val Glu Leu Pro Ile Met Ile Ser Gly Thr Ile Thr Asp 210 215 220	672
gcc tcc ggc cgc acc ctg tgc ggc cag acc acc gag gcc ttc tgg aac Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Thr Glu Ala Phe Trp Asn 225 230 235 240	720
tgc gtg cgg cat gcc cgg ccg atc tgc gta ggc ctg aac tgc gcc ctc Ser Val Arg His Ala Arg Pro Ile Ser Val Gly Leu Asn Cys Ala Leu 245 250 255	768
ggc gcc aag gaa ttg cgg ccg tac atc gag gaa ctg tgc acc aag gcc Gly Ala Lys Glu Leu Arg Pro Tyr Ile Glu Glu Leu Ser Thr Lys Ala 260 265 270	816
gac act cat gtc tgc gcc cac ccc aac gcc ggc ctg ccg aac gcc ttc Asp Thr His Val Ser Ala His Pro Asn Ala Gly Leu Pro Asn Ala Phe 275 280 285	864
ggc gaa tac gac gaa tgc ccg gcg gaa atg gcc gtg gtg gtc gag gaa Gly Glu Tyr Asp Glu Ser Pro Ala Glu Met Ala Val Val Val Glu Glu 290 295 300	912
ttc gcc gcc gcc ggc ttc ctc aat atc gtc ggc ggc tgc tgc ggc acc	960

117																		
Phe	Ala	Ala	Ala	Gly	Phe	Leu	Asn	Ile	Val	Gly	Gly	Cys	Cys	Gly	Thr			
305						310				315					320			
acc	ccg	gcg	cac	atc	gag	gcg	atc	gcc	aag	gca	gtg	gcc	aag	tac	ccg	1008		
Thr	Pro	Ala	His	Ile	Glu	Ala	Ile	Ala	Lys	Ala	Val	Ala	Lys	Tyr	Pro			
				325					330					335				
ccg	cgg	gcc	atc	ccg	gag	att	ccc	cgg	gcc	tgt	cgc	ctg	tcc	ggc	ctg	1056		
Pro	Arg	Ala	Ile	Pro	Glu	Ile	Pro	Arg	Ala	Cys	Arg	Leu	Ser	Gly	Leu			
				340				345					350					
gag	ccg	ttc	acc	atc	gac	cgc	agc	tcg	ctg	ttc	gtc	aac	gtc	ggc	gag	1104		
Glu	Pro	Phe	Thr	Ile	Asp	Arg	Ser	Ser	Leu	Phe	Val	Asn	Val	Gly	Glu			
				355				360				365						
cgc	acc	aac	atc	acc	ggt	tcg	gcc	aag	ttc	gcc	cgg	ctg	atc	cgc	gag	1152		
Arg	Thr	Asn	Ile	Thr	Gly	Ser	Ala	Lys	Phe	Ala	Arg	Leu	Ile	Arg	Glu			
				370			375					380						
gaa	aac	tac	gcg	gaa	gct	ctc	gag	gtc	gcc	cag	cag	cag	gtg	gaa	gcc	1200		
Glu	Asn	Tyr	Ala	Glu	Ala	Leu	Glu	Val	Ala	Gln	Gln	Gln	Val	Glu	Ala			
					390					395					400			
ggc	gcc	cag	gtg	atc	gac	atc	aac	atg	gac	gaa	ggc	atg	ctg	gac	tcg	1248		
Gly	Ala	Gln	Val	Ile	Asp	Ile	Asn	Met	Asp	Glu	Gly	Met	Leu	Asp	Ser			
				405					410					415				
aag	gcg	gcc	atg	gtc	acc	ttc	ctc	aac	ctg	atc	gcc	tcc	gag	ccc	gac	1296		
Lys	Ala	Ala	Met	Val	Thr	Phe	Leu	Asn	Leu	Ile	Ala	Ser	Glu	Pro	Asp			
				420				425					430					
atc	tcg	cgc	gtg	ccg	atc	atg	atc	gac	tcc	tcc	aag	tgg	gaa	gtg	atc	1344		
Ile	Ser	Arg	Val	Pro	Ile	Met	Ile	Asp	Ser	Ser	Lys	Trp	Glu	Val	Ile			
				435			440					445						
gag	gcc	ggc	ctg	aag	tgc	atc	cag	ggc	aag	ggc	atc	gtc	aac	tcg	atc	1392		
Glu	Ala	Gly	Leu	Lys	Cys	Ile	Gln	Gly	Lys	Gly	Ile	Val	Asn	Ser	Ile			
				450		455					460							
tcg	atg	aag	gaa	ggc	gtc	gag	gcc	ttc	aag	cac	cat	gcc	cgc	ctg	tgc	1440		
Ser	Met	Lys	Glu	Gly	Val	Glu	Ala	Phe	Lys	His	His	Ala	Arg	Leu	Cys			
					470					475					480			
aag	cgc	tac	ggc	gcc	gcg	gtg	gtg	gtg	atg	gcc	ttc	gac	gag	gac	ggc	1488		
Lys	Arg	Tyr	Gly	Ala	Ala	Val	Val	Val	Met	Ala	Phe	Asp	Glu	Asp	Gly			
				485					490					495				
cag	gcc	gac	acc	cag	gcg	cgc	aag	gaa	gaa	atc	tgc	aag	cgc	tcc	tac	1536		
Gln	Ala	Asp	Thr	Gln														

ccc tac gcc ctg agc tcg ggc ggg gtg tcc aac gtg tcc ttc tcg ttc	1728
Pro Tyr Ala Leu Ser Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe	
565 570 575	
cgc ggc aac aac ccg gta cgc gag gcg atc cac tcg gtg ttc ctc tac	1776
Arg Gly Asn Asn Pro Val Arg Glu Ala Ile His Ser Val Phe Leu Tyr	
580 585 590	
tac gcg atc cgc aac ggc ctg acc atg ggc atc gtc aac gcc ggc cag	1824
Tyr Ala Ile Arg Asn Gly Leu Thr Met Gly Ile Val Asn Ala Gly Gln	
595 600 605	
ctg gaa atc tac gac gag att ccg aaa gcg ctg cgc gac cgg gtc gag	1872
Leu Glu Ile Tyr Asp Glu Ile Pro Lys Ala Leu Arg Asp Arg Val Glu	
610 615 620	
gac gtg gtg ctc aac cgc acg ccc gag gcc acc gag gcc ctg ctg gcg	1920
Asp Val Val Leu Asn Arg Thr Pro Glu Ala Thr Glu Ala Leu Leu Ala	
625 630 635 640	
atc gcc gac gac tac aag ggc ggc ggc gcg gtc aag gag gcc gag gac	1968
Ile Ala Asp Asp Tyr Lys Gly Gly Gly Ala Val Lys Glu Ala Glu Asp	
645 650 655	
gag gaa tgg cgc agc tac agc gtc gag aag cgc ctc gag cat gcg ctg	2016
Glu Glu Trp Arg Ser Tyr Ser Val Glu Lys Arg Leu Glu His Ala Leu	
660 665 670	
gtc aag ggc atc acc acc tgg atc gtc gag gac acc gag gaa tgc cgc	2064
Val Lys Gly Ile Thr Thr Trp Ile Val Glu Asp Thr Glu Glu Cys Arg	
675 680 685	
cag cag tgt gcg cgt ccc atc gag gtc atc gaa ggt ccg ctg atg tcc	2112
Gln Gln Cys Ala Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Ser	
690 695 700	
ggg atg aac gtg gtc ggc gac ctg ttc ggc gcc ggc aag atg ttc ctc	2160
Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu	
705 710 715 720	
ccg cag gtg gtc aag tcc gcg cga gtg atg aag cag gcg gtg gcc cac	2208
Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His	
725 730 735	
ctg att ccc ttc atc gag gcg gag aaa ggc gac aag ccg gaa gcc aag	2256
Leu Ile Pro Phe Ile Glu Ala Glu Lys Gly Asp Lys Pro Glu Ala Lys	
740 745 750	
ggc aag atc ctg atg gcc acg gtg aag ggc gac gtg cac gac atc ggc	2304
Gly Lys Ile Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly	
755 760 765	
aag aac atc gtc ggc gtg gtg ctc ggc tgc aac ggc tat gac gtg gtc	2352
Lys Asn Ile Val Gly Val Val Leu Gly Cys Asn Gly Tyr Asp Val Val	
770 775 780	
gac ctc ggc gtg atg gtg ccg gcg gag aag atc ctg cag acc gcc atc	2400
Asp Leu Gly Val Met Val Pro Ala Glu Lys Ile Leu Gln Thr Ala Ile	
785 790 795 800	
gcc gag aaa tgc gac atc atc ggc ctg tct ggc ctg atc acg ccg tcg	2448

119

Ala Glu Lys Cys Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser	
805 810 815	
ctg gac gag atg gtc cac gtc gcc aag gaa atg cag cgg cag aat ttc	2496
Leu Asp Glu Met Val His Val Ala Lys Glu Met Gln Arg Gln Asn Phe	
820 825 830	
cag ttg ccg ctg atg atc ggc ggc gcc act acc tcg aag gcg cat acc	2544
Gln Leu Pro Leu Met Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr	
835 840 845	
gcg gtg aag atc gat ccg cag tac agc aac gac gcg gtg gtc tac gtc	2592
Ala Val Lys Ile Asp Pro Gln Tyr Ser Asn Asp Ala Val Val Tyr Val	
850 855 860	
acc gac gcc tcg cgc gcg gta ggc gtg gcc acc agc ctg ctg tcc aag	2640
Thr Asp Ala Ser Arg Ala Val Gly Val Ala Thr Ser Leu Leu Ser Lys	
865 870 875 880	
gag ctg aag gcc gac tac gtg gcc cgc acc cgc gcc gac tac gcg gtg	2688
Glu Leu Lys Ala Asp Tyr Val Ala Arg Thr Arg Ala Asp Tyr Ala Val	
885 890 895	
gtc cgc gaa cgc acg gcc aac cgc agc gcc cgc acc gag cgg ctg agc	2736
Val Arg Glu Arg Thr Ala Asn Arg Ser Ala Arg Thr Glu Arg Leu Ser	
900 905 910	
tac gaa cag gcg atc gcc aac aag ccg gcg ttc gac tgg gcc ggc tac	2784
Tyr Glu Gln Ala Ile Ala Asn Lys Pro Ala Phe Asp Trp Ala Gly Tyr	
915 920 925	
cag gcg ccg acg cct tcc ttc acc ggc gtc agg gtg ctc gac gag atc	2832
Gln Ala Pro Thr Pro Ser Phe Thr Gly Val Arg Val Leu Asp Glu Ile	
930 935 940	
gac ctc gcg gtg ctc gcc gag tac atc gac tgg acg ccg ttc ttc att	2880
Asp Leu Ala Val Leu Ala Glu Tyr Ile Asp Trp Thr Pro Phe Phe Ile	
945 950 955 960	
tcc tgg gac ctg gcc ggc aag tac ccg cgc atc ctc acc gac gag gtg	2928
Ser Trp Asp Leu Ala Gly Lys Tyr Pro Arg Ile Leu Thr Asp Glu Val	
965 970 975	
gtc ggc gag gcc gcc acc tcg ttg ttc aac gac gcc cag gcg atg ctg	2976
Val Gly Glu Ala Ala Thr Ser Leu Phe Asn Asp Ala Gln Ala Met Leu	
980 985 990	
aag aag ctg atc gac gag aag ctg atc aag gcc cgc gcg gtg ttc ggc	3024
Lys Lys Leu Ile Asp Glu Lys Leu Ile Lys Ala Arg Ala Val Phe Gly	
995 1000 1005	
ttc tgg ccg gcc aac cag gtc gag cac gac gac ctg gag gtc tac ggc	3072
Phe Trp Pro Ala Asn Gln Val Glu His Asp Asp Leu Glu Val Tyr Gly	
1010 1015 1020	
gcc gat ggc gag acc ctc gcc acc ctg cac cac ctg ccg cag cag acg	3120
Ala Asp Gly Glu Thr Leu Ala Thr Leu His His Leu Arg Gln Gln Thr	
1025 1030 1035 1040	
atc aag ccg gac ggc aag ccg aac ctg tcg ctg gcc gat ttc gtc gcg	3168
Ile Lys Pro Asp Gly Lys Pro Asn Leu Ser Leu Ala Asp Phe Val Ala	
1045 1050 1055	

120

ccg aag gaa agc ggc gtg cgc gac tac atc ggc ggc ttc atc acc acc 3216  
 Pro Lys Glu Ser Gly Val Arg Asp Tyr Ile Gly Gly Phe Ile Thr Thr  
 1060 1065 1070  
 gcc ggg atc ggc gcc gag gaa gtg gcc aag gcg tac gaa gcc aag ggc 3264  
 Ala Gly Ile Gly Ala Glu Glu Val Ala Lys Ala Tyr Glu Ala Lys Gly  
 1075 1080 1085  
 gac gac tac aac agc atc atg gtc aag gcg ctc gcc gac cgc ctc gcc 3312  
 Asp Asp Tyr Asn Ser Ile Met Val Lys Ala Leu Ala Asp Arg Leu Ala  
 1090 1095 1100  
 gaa gcc tgc gcc gag tgg ctg cac gag cgg gtg cgc aag gag tac tgg 3360  
 Glu Ala Cys Ala Glu Trp Leu His Glu Arg Val Arg Lys Glu Tyr Trp  
 1105 1110 1115 1120  
 ggc tac gcc cgc gac gaa cac ctc gac aac gag gcc ttg atc aag gag 3408  
 Gly Tyr Ala Arg Asp Glu His Leu Asp Asn Glu Ala Leu Ile Lys Glu  
 1125 1130 1135  
 caa tac gtc ggc atc cgc ccg gca ccg ggc tac ccg gcc tgc ccc gac 3456  
 Gln Tyr Val Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp  
 1140 1145 1150  
 cat acc gag aaa ggc act ctg ttc gaa ctg ctc gat ccg cag ggc ctg 3504  
 His Thr Glu Lys Gly Thr Leu Phe Glu Leu Leu Asp Pro Gln Gly Leu  
 1155 1160 1165  
 tcc ggc gtc agc ctg acc gag cac tac gcg atg ttc ccg gcc gcg gcg 3552  
 Ser Gly Val Ser Leu Thr Glu His Tyr Ala Met Phe Pro Ala Ala Ala  
 1170 1175 1180  
 gtc agc ggt tgg tat ttc gcc cac ccg cag gcg cag tac ttc gcg gtc 3600  
 Val Ser Gly Trp Tyr Phe Ala His Pro Gln Ala Gln Tyr Phe Ala Val  
 1185 1190 1195 1200  
 ggc aag atc gac aag gac cag gtg gaa cgc tac agc cag cgc aag ggc 3648  
 Gly Lys Ile Asp Lys Asp Gln Val Glu Arg Tyr Ser Gln Arg Lys Gly  
 1205 1210 1215  
 cag gaa gcc agc gtc agc gag cgc tgg ctg gcg ccg aac ctt ggc tac 3696  
 Gln Glu Ala Ser Val Ser Glu Arg Trp Leu Ala Pro Asn Leu Gly Tyr  
 1220 1225 1230  
 gat gac tga 3705  
 Asp Asp

&lt;210&gt; 28

&lt;211&gt; 1234

&lt;212&gt; PRT

&lt;213&gt; Pseudomonas aeruginosa

&lt;400&gt; 28

Met Ser Ser Pro Leu Thr Asp Arg Ser Ala Arg Leu Gln Ala Leu Gln  
 1 5 10 15

His Ala Leu Arg Glu Arg Ile Leu Ile Leu Asp Gly Gly Met Gly Thr  
 20 25 30

121

Met Ile Gln Ser Tyr Lys Leu Glu Glu Ala Asp Tyr Arg Gly Glu Arg  
 35 40 45  
 Phe Ala Asp Trp Pro Ser Asp Val Lys Gly Asn Asn Asp Leu Leu Leu  
 50 55 60  
 Leu Ser Arg Pro Asp Val Ile Gln Ala Ile Glu Lys Ala Tyr Leu Asp  
 65 70 75 80  
 Ala Gly Ala Asp Ile Leu Glu Thr Asn Thr Phe Asn Ala Thr Gln Val  
 85 90 95  
 Ser Gln Ala Asp Tyr Gly Met Gln Ser Leu Ala Tyr Glu Leu Asn Val  
 100 105 110  
 Glu Gly Ala Arg Leu Ala Arg Gln Val Ala Asp Ala Lys Thr Ala Glu  
 115 120 125  
 Thr Pro Asp Lys Pro Arg Phe Val Ala Gly Val Leu Gly Pro Thr Ser  
 130 135 140  
 Arg Thr Cys Ser Ile Ser Pro Asp Val Asn Asn Pro Gly Tyr Arg Asn  
 145 150 155 160  
 Val Thr Phe Asp Glu Leu Val Glu Asn Tyr Val Glu Ala Thr Arg Gly  
 165 170 175  
 Leu Ile Glu Gly Gly Ala Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp  
 180 185 190  
 Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Val Gln Gly Val Phe Glu  
 195 200 205  
 Glu Leu Gly Val Glu Leu Pro Ile Met Ile Ser Gly Thr Ile Thr Asp  
 210 215 220  
 Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Thr Glu Ala Phe Trp Asn  
 225 230 235 240  
 Ser Val Arg His Ala Arg Pro Ile Ser Val Gly Leu Asn Cys Ala Leu  
 245 250 255  
 Gly Ala Lys Glu Leu Arg Pro Tyr Ile Glu Glu Leu Ser Thr Lys Ala  
 260 265 270  
 Asp Thr His Val Ser Ala His Pro Asn Ala Gly Leu Pro Asn Ala Phe  
 275 280 285  
 Gly Glu Tyr Asp Glu Ser Pro Ala Glu Met Ala Val Val Val Glu Glu  
 290 295 300  
 Phe Ala Ala Ala Gly Phe Leu Asn Ile Val Gly Gly Cys Cys Gly Thr  
 305 310 315 320  
 Thr Pro Ala His Ile Glu Ala Ile Ala Lys Ala Val Ala Lys Tyr Pro  
 325 330 335  
 Pro Arg Ala Ile Pro Glu Ile Pro Arg Ala Cys Arg Leu Ser Gly Leu  
 340 345 350  
 Glu Pro Phe Thr Ile Asp Arg Ser Ser Leu Phe Val Asn Val Gly Glu  
 355 360 365

122

Arg Thr Asn Ile Thr Gly Ser Ala Lys Phe Ala Arg Leu Ile Arg Glu  
 370 375 380  
 Glu Asn Tyr Ala Glu Ala Leu Glu Val Ala Gln Gln Gln Val Glu Ala  
 385 390 395 400  
 Gly Ala Gln Val Ile Asp Ile Asn Met Asp Glu Gly Met Leu Asp Ser  
 405 410 415  
 Lys Ala Ala Met Val Thr Phe Leu Asn Leu Ile Ala Ser Glu Pro Asp  
 420 425 430  
 Ile Ser Arg Val Pro Ile Met Ile Asp Ser Ser Lys Trp Glu Val Ile  
 435 440 445  
 Glu Ala Gly Leu Lys Cys Ile Gln Gly Lys Gly Ile Val Asn Ser Ile  
 450 455 460  
 Ser Met Lys Glu Gly Val Glu Ala Phe Lys His His Ala Arg Leu Cys  
 465 470 475 480  
 Lys Arg Tyr Gly Ala Ala Val Val Val Met Ala Phe Asp Glu Asp Gly  
 485 490 495  
 Gln Ala Asp Thr Gln Ala Arg Lys Glu Glu Ile Cys Lys Arg Ser Tyr  
 500 505 510  
 Asp Ile Leu Val Asp Glu Val Gly Phe Pro Pro Glu Asp Ile Ile Phe  
 515 520 525  
 Asp Ala Asn Ile Phe Ala Ile Ala Thr Gly Ile Glu Glu His Asn Asn  
 530 535 540  
 Tyr Ala Val Asp Phe Ile Asn Ala Cys Ala Tyr Ile Arg Asp Asn Leu  
 545 550 555 560  
 Pro Tyr Ala Leu Ser Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe  
 565 570 575  
 Arg Gly Asn Asn Pro Val Arg Glu Ala Ile His Ser Val Phe Leu Tyr  
 580 585 590  
 Tyr Ala Ile Arg Asn Gly Leu Thr Met Gly Ile Val Asn Ala Gly Gln  
 595 600 605  
 Leu Glu Ile Tyr Asp Glu Ile Pro Lys Ala Leu Arg Asp Arg Val Glu  
 610 615 620  
 Asp Val Val Leu Asn Arg Thr Pro Glu Ala Thr Glu Ala Leu Leu Ala  
 625 630 635 640  
 Ile Ala Asp Asp Tyr Lys Gly Gly Gly Ala Val Lys Glu Ala Glu Asp  
 645 650 655  
 Glu Glu Trp Arg Ser Tyr Ser Val Glu Lys Arg Leu Glu His Ala Leu  
 660 665 670  
 Val Lys Gly Ile Thr Thr Trp Ile Val Glu Asp Thr Glu Glu Cys Arg  
 675 680 685  
 Gln Gln Cys Ala Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Ser



123

690                      695                      700  
 Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu  
 705                      710                      715                      720  
 Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His  
 725                      730                      735  
 Leu Ile Pro Phe Ile Glu Ala Glu Lys Gly Asp Lys Pro Glu Ala Lys  
 740                      745                      750  
 Gly Lys Ile Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly  
 755                      760                      765  
 Lys Asn Ile Val Gly Val Val Leu Gly Cys Asn Gly Tyr Asp Val Val  
 770                      775                      780  
 Asp Leu Gly Val Met Val Pro Ala Glu Lys Ile Leu Gln Thr Ala Ile  
 785                      790                      795                      800  
 Ala Glu Lys Cys Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser  
 805                      810                      815  
 Leu Asp Glu Met Val His Val Ala Lys Glu Met Gln Arg Gln Asn Phe  
 820                      825                      830  
 Gln Leu Pro Leu Met Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr  
 835                      840                      845  
 Ala Val Lys Ile Asp Pro Gln Tyr Ser Asn Asp Ala Val Val Tyr Val  
 850                      855                      860  
 Thr Asp Ala Ser Arg Ala Val Gly Val Ala Thr Ser Leu Leu Ser Lys  
 865                      870                      875                      880  
 Glu Leu Lys Ala Asp Tyr Val Ala Arg Thr Arg Ala Asp Tyr Ala Val  
 885                      890                      895  
 Val Arg Glu Arg Thr Ala Asn Arg Ser Ala Arg Thr Glu Arg Leu Ser  
 900                      905                      910  
 Tyr Glu Gln Ala Ile Ala Asn Lys Pro Ala Phe Asp Trp Ala Gly Tyr  
 915                      920                      925  
 Gln Ala Pro Thr Pro Ser Phe Thr Gly Val Arg Val Leu Asp Glu Ile  
 930                      935                      940  
 Asp Leu Ala Val Leu Ala Glu Tyr Ile Asp Trp Thr Pro Phe Phe Ile  
 945                      950                      955                      960  
 Ser Trp Asp Leu Ala Gly Lys Tyr Pro Arg Ile Leu Thr Asp Glu Val  
 965                      970                      975  
 Val Gly Glu Ala Ala Thr Ser Leu Phe Asn Asp Ala Gln Ala Met Leu  
 980                      985                      990  
 Lys Lys Leu Ile Asp Glu Lys Leu Ile Lys Ala Arg Ala Val Phe Gly  
 995                      1000                      1005  
 Phe Trp Pro Ala Asn Gln Val Glu His Asp Asp Leu Glu Val Tyr Gly  
 1010                      1015                      1020

124

Ala Asp Gly Glu Thr Leu Ala Thr Leu His His Leu Arg Gln Gln Thr  
 1025 1030 1035 1040

Ile Lys Pro Asp Gly Lys Pro Asn Leu Ser Leu Ala Asp Phe Val Ala  
 1045 1050 1055

Pro Lys Glu Ser Gly Val Arg Asp Tyr Ile Gly Gly Phe Ile Thr Thr  
 1060 1065 1070

Ala Gly Ile Gly Ala Glu Glu Val Ala Lys Ala Tyr Glu Ala Lys Gly  
 1075 1080 1085

Asp Asp Tyr Asn Ser Ile Met Val Lys Ala Leu Ala Asp Arg Leu Ala  
 1090 1095 1100

Glu Ala Cys Ala Glu Trp Leu His Glu Arg Val Arg Lys Glu Tyr Trp  
 1105 1110 1115 1120

Gly Tyr Ala Arg Asp Glu His Leu Asp Asn Glu Ala Leu Ile Lys Glu  
 1125 1130 1135

Gln Tyr Val Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp  
 1140 1145 1150

His Thr Glu Lys Gly Thr Leu Phe Glu Leu Leu Asp Pro Gln Gly Leu  
 1155 1160 1165

Ser Gly Val Ser Leu Thr Glu His Tyr Ala Met Phe Pro Ala Ala Ala  
 1170 1175 1180

Val Ser Gly Trp Tyr Phe Ala His Pro Gln Ala Gln Tyr Phe Ala Val  
 1185 1190 1195 1200

Gly Lys Ile Asp Lys Asp Gln Val Glu Arg Tyr Ser Gln Arg Lys Gly  
 1205 1210 1215

Gln Glu Ala Ser Val Ser Glu Arg Trp Leu Ala Pro Asn Leu Gly Tyr  
 1220 1225 1230

Asp Asp

<210> 29  
 <211> 3714  
 <212> DNA  
 <213> Nitrosomas europeae

<220>  
 <221> CDS  
 <222> (1) .. (3711)  
 <223> RNE01732

<400> 29  
 atg aca atg cat gaa cgt gct gat ttg ctg aaa cgg ttg ctt gcc gag 48  
 Met Thr Met His Glu Arg Ala Asp Leu Leu Lys Arg Leu Leu Ala Glu  
 1 5 10 15

cgt atc ctg atg ctc gac ggt gcc atg ggt acg atg atc cag agc tac 96  
 Arg Ile Leu Met Leu Asp Gly Ala Met Gly Thr Met Ile Gln Ser Tyr  
 20 25 30

125

aaa ctg acc gag tcg gat tat cgg ggg gaa cgt ttt gcc gat ttt ccg	144
Lys Leu Thr Glu Ser Asp Tyr Arg Gly Glu Arg Phe Ala Asp Phe Pro	
35 40 45	
cat gat ctc aaa ggc aac aat gat ctg ctc tgc ctg acc aga ccg gaa	192
His Asp Leu Lys Gly Asn Asn Asp Leu Leu Cys Leu Thr Arg Pro Glu	
50 55 60	
gtc atc cgc tcc att cat cgt gct tac ctc gaa gcc ggg tcg gat atc	240
Val Ile Arg Ser Ile His Arg Ala Tyr Leu Glu Ala Gly Ser Asp Ile	
65 70 75 80	
atc gag acc aac acg ttc aac tcg aat gcg ccg tcg atg gcg gac tac	288
Ile Glu Thr Asn Thr Phe Asn Ser Asn Ala Pro Ser Met Ala Asp Tyr	
85 90 95	
cac atg cag gat ctg gtg tat gaa ctg aat gtg gcg ggt gcg cgc ctg	336
His Met Gln Asp Leu Val Tyr Glu Leu Asn Val Ala Gly Ala Arg Leu	
100 105 110	
gcg tgt gag gaa gcg cgg gca atg gaa acg cag caa cct gac cgg ccc	384
Ala Cys Glu Glu Ala Arg Ala Met Glu Thr Gln Gln Pro Asp Arg Pro	
115 120 125	
cgt ttc gtt gcc ggt gtg atc ggg cct acc acc aaa acg gct tca ctc	432
Arg Phe Val Ala Gly Val Ile Gly Pro Thr Thr Lys Thr Ala Ser Leu	
130 135 140	
tca ccg gat gtc aat gat cct gga ttc cgg gcc att acc ttc gat gat	480
Ser Pro Asp Val Asn Asp Pro Gly Phe Arg Ala Ile Thr Phe Asp Asp	
145 150 155 160	
ctg gtg gaa agc tat acc gag tcg gtg cgc ggg ctg atc gac gga ggc	528
Leu Val Glu Ser Tyr Thr Glu Ser Val Arg Gly Leu Ile Asp Gly Gly	
165 170 175	
gcg gat att ctg ctg gtc gaa acc att ttt gac acc ttg aat gcc aaa	576
Ala Asp Ile Leu Leu Val Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys	
180 185 190	
gcc gca ttg ttt gcc atc gat cag tat ttc gaa acg cat gga tta cgt	624
Ala Ala Leu Phe Ala Ile Asp Gln Tyr Phe Glu Thr His Gly Leu Arg	
195 200 205	
ctg ccg gtg atg ata tcg gtc acg att acc gat gct tcg gga cgt aat	672
Leu Pro Val Met Ile Ser Val Thr Ile Thr Asp Ala Ser Gly Arg Asn	
210 215 220	
ctt tcc ggg cag aca ccg gaa gct ttc tgg aat tcg gta cgg cat gca	720
Leu Ser Gly Gln Thr Pro Glu Ala Phe Trp Asn Ser Val Arg His Ala	
225 230 235 240	
cgt ccg ctt tcg gtg gga atc aac tgc gcg ttg ggt gcg gag ttg atg	768
Arg Pro Leu Ser Val Gly Ile Asn Cys Ala Leu Gly Ala Glu Leu Met	
245 250 255	
cgc ccc tac gtg gaa gag ttg tcc aat gtg gct gag gtt ttc acc agc	816
Arg Pro Tyr Val Glu Glu Leu Ser Asn Val Ala Glu Val Phe Thr Ser	
260 265 270	
gcc cat ccc aat gcc ggc ttg cct aat ccc ttg gcg gaa acc ggt tat	864
Ala His Pro Asn Ala Gly Leu Pro Asn Pro Leu Ala Glu Thr Gly Tyr	

126

275	280	285	
gac gaa acg ccg gaa tat acc gcc cgt ctg atc aag gat ttt gcg caa Asp Glu Thr Pro Glu Tyr Thr Ala Arg Leu Ile Lys Asp Phe Ala Gln 290 295 300			912
tcc ggg ttc gtc aac att gtc gcc gcc tgc tgt gcc act aca ccg aaa Ser Gly Phe Val Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Lys 305 310 315 320			960
cat atc gcg gcc att gca gaa gcg gta cgg gac atc cct ccg cgc cca His Ile Ala Ala Ile Ala Glu Ala Val Arg Asp Ile Pro Pro Arg Pro 325 330 335			1008
ctg ccc gat att cct aaa aaa ctg agg ctt tcc gcc ctc gag ccg ctc Leu Pro Asp Ile Pro Lys Lys Leu Arg Leu Ser Gly Leu Glu Pro Leu 340 345 350			1056
aat atc gat gaa cat tcc ctg ttc gta aac gtg ggt gaa cgt acc aat Asn Ile Asp Glu His Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn 355 360 365			1104
gtc acc gcc tcc aag gca ttt gcc cgg ctg att ctc aat gcc ggt tat Val Thr Gly Ser Lys Ala Phe Ala Arg Leu Ile Leu Asn Gly Gly Tyr 370 375 380			1152
gct gaa ggg ctg gtg atc gcg cgc agc cag gtg gag aac gcc gca caa Ala Glu Gly Leu Val Ile Ala Arg Ser Gln Val Glu Asn Gly Ala Gln 385 390 395 400			1200
atc atc gat atc aac atg gat gaa gcg atg ctg gat tca cag aag gcg Ile Ile Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Gln Lys Ala 405 410 415			1248
atg gtg acc ttt ctg aat ctg ctc gct gcc gaa ccg gat atc agc ccg Met Val Thr Phe Leu Asn Leu Leu Ala Ala Glu Pro Asp Ile Ser Arg 420 425 430			1296
ctg ccg atc atg ctc gat tcc agc aaa tgg tgc gtg atc gaa gcc gga Leu Pro Ile Met Leu Asp Ser Ser Lys Trp Ser Val Ile Glu Ala Gly 435 440 445			1344
ctg aaa tgt gtc cag ggt aag gcg gtc atc aat tcc atc agc ctc aag Leu Lys Cys Val Gln Gly Lys Ala Val Ile Asn Ser Ile Ser Leu Lys 450 455 460			1392
gaa ggt gaa gcg gag ttt tta cat cat gcc agg ctg gcg cgt cgt tat Glu Gly Glu Ala Glu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr 465 470 475 480			1440
ggg gcc gcg gtg att gtc atg gct ttc gac gaa acc ggg cag gcc gat Gly Ala Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp 485 490 495			1488
acc ttg cag cgc aag gtg gaa atc tgc acg cgt tgt tac cat aca ctg Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu 500 505 510			1536
att gaa cag gcc gat ttc cca ccc gag gat atc att ttc gac ccc aat Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn 515 520 525			1584

127

att ttt gcc att gct acg ggt atc gaa gaa cac agt aac tat gca gtg Ile Phe Ala Ile Ala Thr Gly Ile Glu Glu His Ser Asn Tyr Ala Val 530 535 540	1632
gat ttt atc gag gcg aca cac gtc atc cgg caa acg ctg cct tat gcc Asp Phe Ile Glu Ala Thr His Val Ile Arg Gln Thr Leu Pro Tyr Ala 545 550 555 560	1680
aaa gtc agc ggg ggt gtt tcc aat gtt tcc ttc tcg ttc cgg ggt aac Lys Val Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn 565 570 575	1728
gaa ccg atc cgc gaa gcc att cat acc gca ttc ctg tat cac gcg gtc Glu Pro Ile Arg Glu Ala Ile His Thr Ala Phe Leu Tyr His Ala Val 580 585 590	1776
aag gca ggc atg acc atg ggt atc gtc aac gca ggt cag ctt ggg gtt Lys Ala Gly Met Thr Met Gly Ile Val Asn Ala Gly Gln Leu Gly Val 595 600 605	1824
tat tcc gac att ccg ccc gat ctg ctg gaa cat gtc gag gat gta ctg Tyr Ser Asp Ile Pro Pro Asp Leu Leu Glu His Val Glu Asp Val Leu 610 615 620	1872
ctg aac cgg cgg cct gat gca acc gaa cgt ctg gtg gag ttt gcg gaa Leu Asn Arg Arg Pro Asp Ala Thr Glu Arg Leu Val Glu Phe Ala Glu 625 630 635 640	1920
cat ttc aag gga cag aaa aag gag cag atc gaa gat ctg tcc tgg cgt His Phe Lys Gly Gln Lys Lys Glu Gln Ile Glu Asp Leu Ser Trp Arg 645 650 655	1968
gat gaa ccg gtg cgg cag cgc ctg att cat gca ctg gtc agg ggt atc Asp Glu Pro Val Arg Gln Arg Leu Ile His Ala Leu Val Arg Gly Ile 660 665 670	2016
agc acc tac atc gtc gag gat acc gag ctc gtc cgg cag gag atc gac Ser Thr Tyr Ile Val Glu Asp Thr Glu Leu Val Arg Gln Glu Ile Asp 675 680 685	2064
agc cag gga ggc aag ccg atc gag gtg atc gaa ggc ccg ctc atg gac Ser Gln Gly Gly Lys Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp 690 695 700	2112
ggc atg aat gta gtg ggg gat ctg ttt ggc gca ggc aag atg ttt ctg Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu 705 710 715 720	2160
cca cag gtg gtc aag tcg gca cgg gtg atg aag cag gcg gtt gcc tat Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr 725 730 735	2208
ctg ttg ccg tac atc gag gca gag aaa aaa att tcc ggc gac agc aag Leu Leu Pro Tyr Ile Glu Ala Glu Lys Lys Ile Ser Gly Asp Ser Lys 740 745 750	2256
ccc aag ggc aag gtg gtg atc gct acc gtc aaa ggg gat gtg cat gat Pro Lys Gly Lys Val Val Ile Ala Thr Val Lys Gly Asp Val His Asp 755 760 765	2304
att ggc aag aat atc gtt tcc gtc gtg ttg cag tgt aat aac ttt gaa Ile Gly Lys Asn Ile Val Ser Val Val Leu Gln Cys Asn Asn Phe Glu	2352

128

770	775	780	
gtc atc aac atg ggg gtg atg gtc ccc agt gca cag att ctg gaa aca Val Ile Asn Met Gly Val Met Val Pro Ser Ala Gln Ile Leu Glu Thr 785 790 795 800			2400
gca cgc cgt gaa cag gtc gat atg atc ggt ctg tcc ggc ctg atc acc Ala Arg Arg Glu Gln Val Asp Met Ile Gly Leu Ser Gly Leu Ile Thr 805 810 815			2448
cct tcg ctg gaa gaa atg gcg cat gtt gcc cgg gaa atg gag cgt gaa Pro Ser Leu Glu Glu Met Ala His Val Ala Arg Glu Met Glu Arg Glu 820 825 830			2496
caa ttc acc gtt ccg ctg ctg atc ggt ggc gcc acc act tcg cgg atg Gln Phe Thr Val Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Met 835 840 845			2544
cat acg gca gtc aaa atc gca ccc cat tac ggt ggg gtg acc gta tgg His Thr Ala Val Lys Ile Ala Pro His Tyr Gly Gly Val Thr Val Trp 850 855 860			2592
gtg ccg gat gcc agc cgg gca gtc ggg gtg tgc agc aat ctg atg tca Val Pro Asp Ala Ser Arg Ala Val Gly Val Cys Ser Asn Leu Met Ser 865 870 875 880			2640
cag gat ctg cgt gat gac tat gtc cgg cag gtc aag gcc gag cag gag Gln Asp Leu Arg Asp Asp Tyr Val Arg Gln Val Lys Ala Glu Gln Glu 885 890 895			2688
aag agc cgg gtg cag cac cgc aac aag aaa ggg cca tcc aag ctc ctc Lys Ser Arg Val Gln His Arg Asn Lys Lys Gly Pro Ser Lys Leu Leu 900 905 910			2736
act ttc gag gaa gcc cgg gcc aac gca ctc aag acg gat tgg gct cgt Thr Phe Glu Glu Ala Arg Ala Asn Ala Leu Lys Thr Asp Trp Ala Arg 915 920 925			2784
tat act cca cca gct ccg gat ttc ctg ggg ttg cgc acc ctc aac aac Tyr Thr Pro Pro Ala Pro Asp Phe Leu Gly Leu Arg Thr Leu Asn Asn 930 935 940			2832
tat ccg ctg gaa aca ctg gtg ccg cac atc gac tgg aca cct ttc ttc Tyr Pro Leu Glu Thr Leu Val Pro His Ile Asp Trp Thr Pro Phe Phe 945 950 955 960			2880
cag gca tgg gaa ctg cac ggg cgc tat cct gcc atc ctg cag gat gaa Gln Ala Trp Glu Leu His Gly Arg Tyr Pro Ala Ile Leu Gln Asp Glu 965 970 975			2928
ctc gtc ggg gaa gca gcc agc aat ctg ttt cgc gat gcc cag aat atg Leu Val Gly Glu Ala Ala Ser Asn Leu Phe Arg Asp Ala Gln Asn Met 980 985 990			2976
ctc aga aaa atc gtc gag caa aaa tgg ctc acc gcc aac gcc gtt atc Leu Arg Lys Ile Val Glu Gln Lys Trp Leu Thr Ala Asn Ala Val Ile 995 1000 1005			3024
ggc ctg ttc ccg gcc aat acc gtc aat gga gat gat atc gag att tat Gly Leu Phe Pro Ala Asn Thr Val Asn Gly Asp Asp Ile Glu Ile Tyr 1010 1015 1020			3072

129

gct gac cgt agt cgc agt cag gtg atc atg acc tgg cac acc ttg cgg 3120  
 Ala Asp Arg Ser Arg Ser Gln Val Ile Met Thr Trp His Thr Leu Arg  
 1025 1030 1035 1040

cag cag acg gcc aaa ccg gca ggg cgt ccc aat ctg gca ctg gct gat 3168  
 Gln Gln Thr Ala Lys Pro Ala Gly Arg Pro Asn Leu Ala Leu Ala Asp  
 1045 1050 1055

ttc att gcg ccg cgt gaa acc gga ctg gac gat acc atc ggt ttg ttt 3216  
 Phe Ile Ala Pro Arg Glu Thr Gly Leu Asp Asp Thr Ile Gly Leu Phe  
 1060 1065 1070

gcc gtc agc gcc ggt ttc ggt atc gat gaa cgc ata cgc gct ttt gaa 3264  
 Ala Val Ser Ala Gly Phe Gly Ile Asp Glu Arg Ile Arg Ala Phe Glu  
 1075 1080 1085

gct gca aac gat gat tac agt gcc atc atc ctg aaa gca ctg gct gat 3312  
 Ala Ala Asn Asp Asp Tyr Ser Ala Ile Ile Leu Lys Ala Leu Ala Asp  
 1090 1095 1100

cgt ctg gct gaa gcg ttt gca gaa cac atg cat gca cgg gtg cgg cga 3360  
 Arg Leu Ala Glu Ala Phe Ala Glu His Met His Ala Arg Val Arg Arg  
 1105 1110 1115 1120

gaa ttc tgg ggc tat gtg aaa gat gag agt ctg gac aat gaa cag ttg 3408  
 Glu Phe Trp Gly Tyr Val Lys Asp Glu Ser Leu Asp Asn Glu Gln Leu  
 1125 1130 1135

atc gac gag caa tac ctg gga atc cgt cca gca cca ggt tat cct gcc 3456  
 Ile Asp Glu Gln Tyr Leu Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala  
 1140 1145 1150

tgc cct gat cat acc gaa aag ggg cca ttg ttc gct ctg ctg gaa gcg 3504  
 Cys Pro Asp His Thr Glu Lys Gly Pro Leu Phe Ala Leu Leu Glu Ala  
 1155 1160 1165

gaa aaa cgc agc gga atc gtc ata acg gaa tca ttt gcc atg gtg ccg 3552  
 Glu Lys Arg Ser Gly Ile Val Ile Thr Glu Ser Phe Ala Met Val Pro  
 1170 1175 1180

act gca gca gta tcc ggc ttc tat ctc tct tac cct gaa tcc agc tat 3600  
 Thr Ala Ala Val Ser Gly Phe Tyr Leu Ser Tyr Pro Glu Ser Ser Tyr  
 1185 1190 1195 1200

ttt gct gtt gga aaa atc gga aaa gat cag gtc gag gat tat gca aga 3648  
 Phe Ala Val Gly Lys Ile Gly Lys Asp Gln Val Glu Asp Tyr Ala Arg  
 1205 1210 1215

cgc aaa ggg tgg acg ctg gaa gaa gca gaa agg tgg ctt gcg cct gtc 3696  
 Arg Lys Gly Trp Thr Leu Glu Glu Ala Glu Arg Trp Leu Ala Pro Val  
 1220 1225 1230

ttg gcg tat gag cgt taa 3714  
 Leu Ala Tyr Glu Arg  
 1235

&lt;210&gt; 30

&lt;211&gt; 1237

&lt;212&gt; PRT

&lt;213&gt; Nitrosomas europeae

130

&lt;400&gt; 30

```

Met Thr Met His Glu Arg Ala Asp Leu Leu Lys Arg Leu Leu Ala Glu
 1           5           10           15

Arg Ile Leu Met Leu Asp Gly Ala Met Gly Thr Met Ile Gln Ser Tyr
      20           25           30

Lys Leu Thr Glu Ser Asp Tyr Arg Gly Glu Arg Phe Ala Asp Phe Pro
      35           40           45

His Asp Leu Lys Gly Asn Asn Asp Leu Leu Cys Leu Thr Arg Pro Glu
      50           55           60

Val Ile Arg Ser Ile His Arg Ala Tyr Leu Glu Ala Gly Ser Asp Ile
      65           70           75           80

Ile Glu Thr Asn Thr Phe Asn Ser Asn Ala Pro Ser Met Ala Asp Tyr
      85           90           95

His Met Gln Asp Leu Val Tyr Glu Leu Asn Val Ala Gly Ala Arg Leu
      100          105          110

Ala Cys Glu Glu Ala Arg Ala Met Glu Thr Gln Gln Pro Asp Arg Pro
      115          120          125

Arg Phe Val Ala Gly Val Ile Gly Pro Thr Thr Lys Thr Ala Ser Leu
      130          135          140

Ser Pro Asp Val Asn Asp Pro Gly Phe Arg Ala Ile Thr Phe Asp Asp
      145          150          155          160

Leu Val Glu Ser Tyr Thr Glu Ser Val Arg Gly Leu Ile Asp Gly Gly
      165          170          175

Ala Asp Ile Leu Leu Val Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys
      180          185          190

Ala Ala Leu Phe Ala Ile Asp Gln Tyr Phe Glu Thr His Gly Leu Arg
      195          200          205

Leu Pro Val Met Ile Ser Val Thr Ile Thr Asp Ala Ser Gly Arg Asn
      210          215          220

Leu Ser Gly Gln Thr Pro Glu Ala Phe Trp Asn Ser Val Arg His Ala
      225          230          235          240

Arg Pro Leu Ser Val Gly Ile Asn Cys Ala Leu Gly Ala Glu Leu Met
      245          250          255

Arg Pro Tyr Val Glu Glu Leu Ser Asn Val Ala Glu Val Phe Thr Ser
      260          265          270

Ala His Pro Asn Ala Gly Leu Pro Asn Pro Leu Ala Glu Thr Gly Tyr
      275          280          285

Asp Glu Thr Pro Glu Tyr Thr Ala Arg Leu Ile Lys Asp Phe Ala Gln
      290          295          300

Ser Gly Phe Val Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Lys
      305          310          315          320

His Ile Ala Ala Ile Ala Glu Ala Val Arg Asp Ile Pro Pro Arg Pro

```



131

325

330

335

Leu Pro Asp Ile Pro Lys Lys Leu Arg Leu Ser Gly Leu Glu Pro Leu  
 340 345 350  
 Asn Ile Asp Glu His Ser Leu Phe Val Asn Val Gly Glu Arg Thr Asn  
 355 360 365  
 Val Thr Gly Ser Lys Ala Phe Ala Arg Leu Ile Leu Asn Gly Gly Tyr  
 370 375 380  
 Ala Glu Gly Leu Val Ile Ala Arg Ser Gln Val Glu Asn Gly Ala Gln  
 385 390 395 400  
 Ile Ile Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Gln Lys Ala  
 405 410 415  
 Met Val Thr Phe Leu Asn Leu Leu Ala Ala Glu Pro Asp Ile Ser Arg  
 420 425 430  
 Leu Pro Ile Met Leu Asp Ser Ser Lys Trp Ser Val Ile Glu Ala Gly  
 435 440 445  
 Leu Lys Cys Val Gln Gly Lys Ala Val Ile Asn Ser Ile Ser Leu Lys  
 450 455 460  
 Glu Gly Glu Ala Glu Phe Leu His His Ala Arg Leu Ala Arg Arg Tyr  
 465 470 475 480  
 Gly Ala Ala Val Ile Val Met Ala Phe Asp Glu Thr Gly Gln Ala Asp  
 485 490 495  
 Thr Leu Gln Arg Lys Val Glu Ile Cys Thr Arg Cys Tyr His Thr Leu  
 500 505 510  
 Ile Glu Gln Ala Asp Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn  
 515 520 525  
 Ile Phe Ala Ile Ala Thr Gly Ile Glu Glu His Ser Asn Tyr Ala Val  
 530 535 540  
 Asp Phe Ile Glu Ala Thr His Val Ile Arg Gln Thr Leu Pro Tyr Ala  
 545 550 555 560  
 Lys Val Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn  
 565 570 575  
 Glu Pro Ile Arg Glu Ala Ile His Thr Ala Phe Leu Tyr His Ala Val  
 580 585 590  
 Lys Ala Gly Met Thr Met Gly Ile Val Asn Ala Gly Gln Leu Gly Val  
 595 600 605  
 Tyr Ser Asp Ile Pro Pro Asp Leu Leu Glu His Val Glu Asp Val Leu  
 610 615 620  
 Leu Asn Arg Arg Pro Asp Ala Thr Glu Arg Leu Val Glu Phe Ala Glu  
 625 630 635 640  
 His Phe Lys Gly Gln Lys Lys Glu Gln Ile Glu Asp Leu Ser Trp Arg  
 645 650 655

132

Asp Glu Pro Val Arg Gln Arg Leu Ile His Ala Leu Val Arg Gly Ile  
 660 665 670  
 Ser Thr Tyr Ile Val Glu Asp Thr Glu Leu Val Arg Gln Glu Ile Asp  
 675 680 685  
 Ser Gln Gly Gly Lys Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp  
 690 695 700  
 Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu  
 705 710 715 720  
 Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala Tyr  
 725 730 735  
 Leu Leu Pro Tyr Ile Glu Ala Glu Lys Lys Ile Ser Gly Asp Ser Lys  
 740 745 750  
 Pro Lys Gly Lys Val Val Ile Ala Thr Val Lys Gly Asp Val His Asp  
 755 760 765  
 Ile Gly Lys Asn Ile Val Ser Val Val Leu Gln Cys Asn Asn Phe Glu  
 770 775 780  
 Val Ile Asn Met Gly Val Met Val Pro Ser Ala Gln Ile Leu Glu Thr  
 785 790 795 800  
 Ala Arg Arg Glu Gln Val Asp Met Ile Gly Leu Ser Gly Leu Ile Thr  
 805 810 815  
 Pro Ser Leu Glu Glu Met Ala His Val Ala Arg Glu Met Glu Arg Glu  
 820 825 830  
 Gln Phe Thr Val Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Met  
 835 840 845  
 His Thr Ala Val Lys Ile Ala Pro His Tyr Gly Gly Val Thr Val Trp  
 850 855 860  
 Val Pro Asp Ala Ser Arg Ala Val Gly Val Cys Ser Asn Leu Met Ser  
 865 870 875 880  
 Gln Asp Leu Arg Asp Asp Tyr Val Arg Gln Val Lys Ala Glu Gln Glu  
 885 890 895  
 Lys Ser Arg Val Gln His Arg Asn Lys Lys Gly Pro Ser Lys Leu Leu  
 900 905 910  
 Thr Phe Glu Glu Ala Arg Ala Asn Ala Leu Lys Thr Asp Trp Ala Arg  
 915 920 925  
 Tyr Thr Pro Pro Ala Pro Asp Phe Leu Gly Leu Arg Thr Leu Asn Asn  
 930 935 940  
 Tyr Pro Leu Glu Thr Leu Val Pro His Ile Asp Trp Thr Pro Phe Phe  
 945 950 955 960  
 Gln Ala Trp Glu Leu His Gly Arg Tyr Pro Ala Ile Leu Gln Asp Glu  
 965 970 975  
 Leu Val Gly Glu Ala Ala Ser Asn Leu Phe Arg Asp Ala Gln Asn Met  
 980 985 990

133

Leu Arg Lys Ile Val Glu Gln Lys Trp Leu Thr Ala Asn Ala Val Ile  
 995 1000 1005  
 Gly Leu Phe Pro Ala Asn Thr Val Asn Gly Asp Asp Ile Glu Ile Tyr  
 1010 1015 1020  
 Ala Asp Arg Ser Arg Ser Gln Val Ile Met Thr Trp His Thr Leu Arg  
 1025 1030 1035 1040  
 Gln Gln Thr Ala Lys Pro Ala Gly Arg Pro Asn Leu Ala Leu Ala Asp  
 1045 1050 1055  
 Phe Ile Ala Pro Arg Glu Thr Gly Leu Asp Asp Thr Ile Gly Leu Phe  
 1060 1065 1070  
 Ala Val Ser Ala Gly Phe Gly Ile Asp Glu Arg Ile Arg Ala Phe Glu  
 1075 1080 1085  
 Ala Ala Asn Asp Asp Tyr Ser Ala Ile Ile Leu Lys Ala Leu Ala Asp  
 1090 1095 1100  
 Arg Leu Ala Glu Ala Phe Ala Glu His Met His Ala Arg Val Arg Arg  
 1105 1110 1115 1120  
 Glu Phe Trp Gly Tyr Val Lys Asp Glu Ser Leu Asp Asn Glu Gln Leu  
 1125 1130 1135  
 Ile Asp Glu Gln Tyr Leu Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala  
 1140 1145 1150  
 Cys Pro Asp His Thr Glu Lys Gly Pro Leu Phe Ala Leu Leu Glu Ala  
 1155 1160 1165  
 Glu Lys Arg Ser Gly Ile Val Ile Thr Glu Ser Phe Ala Met Val Pro  
 1170 1175 1180  
 Thr Ala Ala Val Ser Gly Phe Tyr Leu Ser Tyr Pro Glu Ser Ser Tyr  
 1185 1190 1195 1200  
 Phe Ala Val Gly Lys Ile Gly Lys Asp Gln Val Glu Asp Tyr Ala Arg  
 1205 1210 1215  
 Arg Lys Gly Trp Thr Leu Glu Glu Ala Glu Arg Trp Leu Ala Pro Val  
 1220 1225 1230  
 Leu Ala Tyr Glu Arg  
 1235

<210> 31  
 <211> 3774  
 <212> DNA  
 <213> Bordetella pertussis

<220>  
 <221> CDS  
 <222> (1) .. (3771)  
 <223> RBP00104

<220>  
 <221> unsure

134

&lt;222&gt; 205 .. 205

&lt;223&gt; All occurrences of n indicate any nucleotide

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 277 .. 277

&lt;223&gt; All occurrences of n indicate any nucleotide

&lt;400&gt; 31

```

gtg cct tat ccc cgt atc ccc ttc ccg ctg tcc gcc tac acg cat ggc      48
Val Pro Tyr Pro Arg Ile Pro Phe Pro Leu Ser Ala Tyr Thr His Gly
  1             5             10             15

ggc gag ttc gtc cgc caa ctg gac aag cgc atc ctg atc ctg gat ggt      96
Gly Glu Phe Val Arg Gln Leu Asp Lys Arg Ile Leu Ile Leu Asp Gly
          20             25             30

gcc atg ggc acg atg atc cag cgc tac aag ctg ggc gag gcc gat ttc      144
Ala Met Gly Thr Met Ile Gln Arg Tyr Lys Leu Gly Glu Ala Asp Phe
      35             40             45

cgt ggc gag cgc ttc gcc gag cac cac aag gat ctc aag ggc gac aac      192
Arg Gly Glu Arg Phe Ala Gln His His Lys Asp Leu Lys Gly Asp Asn
  50             55             60

gaa ctg ctg tcg ntg gtg cgc ccg gac gtg atc gcg gaa atc cac cgg      240
Glu Leu Leu Ser Xaa Val Arg Pro Asp Val Ile Ala Glu Ile His Arg
  65             70             75             80

cag tac ctc gag gcc ggc gcc gac gtg atc gag acc nac acc ttc ggc      288
Gln Tyr Leu Glu Ala Gly Ala Asp Val Ile Glu Thr Xaa Thr Phe Gly
          85             90             95

gcc acg tcg atc gcc cag ggc gat tac gac ctg ccg gag ctg gcc tac      336
Ala Thr Ser Ile Ala Gln Gly Asp Tyr Asp Leu Pro Glu Leu Ala Tyr
      100             105             110

gag atg aac ctg gag tcg gcc cgc ctg gcg cgc gcc gcc tgc gac gcc      384
Glu Met Asn Leu Glu Ser Ala Arg Leu Ala Arg Ala Ala Cys Asp Ala
      115             120             125

tac agc acg ccc gag cat ccg cgc ttc gtg gcc ggg gcg ctg ggg ccg      432
Tyr Ser Thr Pro Glu His Pro Arg Phe Val Ala Gly Ala Leu Gly Pro
      130             135             140

cag ccc aag acc gcg tcc atc tcg ccc gac gtc aac gac ccg ggg gcg      480
Gln Pro Lys Thr Ala Ser Ile Ser Pro Asp Val Asn Asp Pro Gly Ala
      145             150             155             160

cgc aac gtc acc ttc gac gag ctg cgc gcg gcc tat gtc gag cag ctc      528
Arg Asn Val Thr Phe Asp Glu Leu Arg Ala Ala Tyr Val Glu Gln Leu
          165             170             175

aat ggc ctg ctc gac ggc ggc atc gac atc gtc ctg atc gaa acc atc      576
Asn Gly Leu Leu Asp Gly Gly Ile Asp Ile Val Leu Ile Glu Thr Ile
      180             185             190

ttc gat acg ctc aac gcc aag gcg gcc atc ttc gcc gtc gag gaa gcg      624
Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Val Glu Glu Ala
      195             200             205

ttc gag gcg cgc ggc gtg cgc ctg ccg gtg atg att tcg ggc acc gtg      672

```

135

Phe	Glu	Ala	Arg	Gly	Val	Arg	Leu	Pro	Val	Met	Ile	Ser	Gly	Thr	Val		
210						215					220						
acc	gat	gcg	tcg	ggc	cgc	atc	ctg	tcc	ggc	cag	acc	gtc	gag	gcg	ttc	720	
Thr	Asp	Ala	Ser	Gly	Arg	Ile	Leu	Ser	Gly	Gln	Thr	Val	Glu	Ala	Phe		
225					230					235					240		
tgg	aac	tcg	gtg	cgc	cat	gcg	cgg	ccg	gtc	acc	atc	ggc	ctg	aac	tgc	768	
Trp	Asn	Ser	Val	Arg	His	Ala	Arg	Pro	Val	Thr	Ile	Gly	Leu	Asn	Cys		
				245					250						255		
gcg	ctg	ggc	gcg	gcg	ctg	atg	cgt	ccg	tat	gtg	gcc	gag	ctg	tcc	aag	816	
Ala	Leu	Gly	Ala	Ala	Leu	Met	Arg	Pro	Tyr	Val	Ala	Glu	Leu	Ser	Lys		
			260					265							270		
atc	tgc	gac	acc	tat	gtg	tgc	gtc	tat	ccc	aac	gcc	ggc	ctg	ccc	aat	864	
Ile	Cys	Asp	Thr	Tyr	Val	Cys	Val	Tyr	Pro	Asn	Ala	Gly	Leu	Pro	Asn		
		275						280							285		
ccc	atg	gcc	gag	acg	ggc	ttt	gac	gaa	acg	ccg	gcc	gat	acc	tcg	gcc	912	
Pro	Met	Ala	Glu	Thr	Gly	Phe	Asp	Glu	Thr	Pro	Ala	Asp	Thr	Ser	Ala		
		290				295									300		
ctg	ctg	gaa	gag	ttc	gcc	cag	gcc	ggg	ctg	gtc	aac	atg	gcc	ggc	ggc	960	
Leu	Leu	Glu	Glu	Phe	Ala	Gln	Ala	Gly	Leu	Val	Asn	Met	Ala	Gly	Gly		
305					310					315					320		
tgt	tgc	ggc	acc	acg	ccc	gag	cac	atc	cgc	gcc	atc	gcc	ggc	aag	gtg	1008	
Cys	Cys	Gly	Thr	Thr	Pro	Glu	His	Ile	Arg	Ala	Ile	Ala	Gly	Lys	Val		
				325					330						335		
gcc	gcg	ctg	acg	ccg	cgc	gcg	gtg	ccc	gag	gtg	ccg	gtc	aag	acc	cgc	1056	
Ala	Ala	Leu	Thr	Pro	Arg	Ala	Val	Pro	Glu	Val	Pro	Val	Lys	Thr	Arg		
			340					345							350		
ctg	tcg	ggc	ctg	gag	gcg	ctc	aac	atc	gac	gac	gag	act	ctg	ttc	gtc	1104	
Leu	Ser	Gly	Leu	Glu	Ala	Leu	Asn	Ile	Asp	Asp	Glu	Thr	Leu	Phe	Val		
		355					360								365		
aac	gtg	ggc	gag	cgc	acc	aac	gtg	acg	ggc	agc	aag	atg	ttc	gcc	cgc	1152	
Asn	Val	Gly	Glu	Arg	Thr	Asn	Val	Thr	Gly	Ser	Lys	Met	Phe	Ala	Arg		
		370					375								380		
ctg	gtc	cgc	gag	gag	aaa	tac	gac	gag	gcg	ctg	gcc	gtg	gcg	cgc	cag	1200	
Leu	Val	Arg	Glu	Glu	Lys	Tyr	Asp	Glu	Ala	Leu	Ala	Val	Ala	Arg	Gln		
385					390					395					400		
cag	gtc	gag	aac	ggg	gcc	cag	atc	atc	gac	gtc	aac	atg	gac	gag	gcg	1248	
Gln	Val	Glu	Asn	Gly	Ala	Gln	Ile	Ile	Asp	Val	Asn	Met	Asp	Glu	Ala		
				405					410						415		
atg	ctg	gac	tcg	gtg	gcc	tgt	atg	cac	cgc	ttc	ctc	aac	ctg	atc	gcg	1296	
Met	Leu	Asp	Ser	Val	Ala	Cys	Met	His	Arg	Phe	Leu	Asn	Leu	Ile	Ala		
			420						425						430		
tcc	gag	ccc	gac	atc	gcg	cgg	gtg	ccg	gtg	atg	atc	gac	agt	tcc	aag	1344	
Ser	Glu	Pro	Asp	Ile	Ala	Arg	Val	Pro	Val	Met	Ile	Asp	Ser	Ser	Lys		
			435					440							445		
tgg	gaa	gtg	atc	gag	acc	ggc	ctg	aag	tgc	gtg	cag	ggc	aag	gcc	gtg	1392	
Trp	Glu	Val	Ile	Glu	Thr	Gly	Leu	Lys	Cys	Val	Gln	Gly	Lys	Ala	Val		
		450					455								460		

gtc aac tcg atc tcc atg aag gaa ggc gag gag ccg ttc cgc cat cat	1440
Val Asn Ser Ile Ser Met Lys Glu Gly Glu Glu Pro Phe Arg His His	
465 470 475 480	
gcg cgc ctg tgc cgc cgc tac ggc gcg gcc atg gtg gtc atg gcc ttc	1488
Ala Arg Leu Cys Arg Arg Tyr Gly Ala Ala Met Val Val Met Ala Phe	
485 490 495	
gac gaa cag ggg cag gcc gac tcg ctg gag cgc cgc aag gaa atc tgc	1536
Asp Glu Gln Gly Gln Ala Asp Ser Leu Glu Arg Arg Lys Glu Ile Cys	
500 505 510	
ggc cgc gcc tac cgt atc ctg gtc gag gaa gag ggc ttc ccg ccc gag	1584
Gly Arg Ala Tyr Arg Ile Leu Val Glu Glu Glu Gly Phe Pro Pro Glu	
515 520 525	
gac atc atc ttc gat ccc aac gtg ttc gcg gtg gcc acc ggc atc gac	1632
Asp Ile Ile Phe Asp Pro Asn Val Phe Ala Val Ala Thr Gly Ile Asp	
530 535 540	
gaa cac aat cac tac gcc gtc gat ttc atc gaa ggc gcg cgc tgg atc	1680
Glu His Asn His Tyr Ala Val Asp Phe Ile Glu Gly Ala Arg Trp Ile	
545 550 555 560	
cgc gcg aac ctg ccg cat gcc cgc att tcg ggc ggc atc tcg aac gtc	1728
Arg Ala Asn Leu Pro His Ala Arg Ile Ser Gly Gly Ile Ser Asn Val	
565 570 575	
agc ttc tcg ttc cgc ggc aac gag ccg atg cgc gag gcg atc cat acc	1776
Ser Phe Ser Phe Arg Gly Asn Glu Pro Met Arg Glu Ala Ile His Thr	
580 585 590	
gtc ttc ctg tac tac gcc atc gag gcc ggc ctg acg atg ggc atc gtc	1824
Val Phe Leu Tyr Tyr Ala Ile Glu Ala Gly Leu Thr Met Gly Ile Val	
595 600 605	
aac gcg ggc cag ctg ggc gta tat gcc gac ctg gcg ccg cac ctg cgc	1872
Asn Ala Gly Gln Leu Gly Val Tyr Ala Asp Leu Ala Pro His Leu Arg	
610 615 620	
gac ctg gtc gag gac gtc atc ctg gac cgc ccc gag ccg gtg ggc cgc	1920
Asp Leu Val Glu Asp Val Ile Leu Asp Arg Pro Glu Pro Val Gly Arg	
625 630 635 640	
agc gac tcg gcc gac gag cgc tcg ccc acc gaa cgg ctg gtg cag ttt	1968
Ser Asp Ser Ala Asp Glu Arg Ser Pro Thr Glu Arg Leu Val Gln Phe	
645 650 655	
gcc gag acc gtc aag ggc tcg ggc gcg aag aag gaa gaa gac ctg acc	2016
Ala Glu Thr Val Lys Gly Ser Gly Ala Lys Lys Glu Glu Asp Leu Thr	
660 665 670	
tgg cgc acc ggc tcg gtc gag cag cgc ctg gcg cat gcc ctg gtg cac	2064
Trp Arg Thr Gly Ser Val Glu Gln Arg Leu Ala His Ala Leu Val His	
675 680 685	
ggc atc acc acc ttc atc gtc gag gac acc gag gaa gtg cgc cag cag	2112
Gly Ile Thr Thr Phe Ile Val Glu Asp Thr Glu Glu Val Arg Gln Gln	
690 695 700	
gtc gcc gcg cgc ggc ggg cgc acc atc gaa gtg atc gaa ggt ccg ctg	2160

	137																								
Val Ala Ala Arg Gly Gly Arg Thr Ile Glu Val Ile Glu Gly Pro Leu 705	710								715								720								
atg gac ggc atg aac gtg gtc ggc gac ctg ttc ggc gcg ggc aag atg Met Asp Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met	725								730								735								2208
ttc ctg ccg caa gtg gtg aag tcg gcg cgc gtg atg aag cag gcg gtg Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val	740								745								750								2256
gcg cac ctg att ccc ttc atc gag gag gaa aag cgc cag atc gcg gcc Ala His Leu Ile Pro Phe Ile Glu Glu Glu Lys Arg Gln Ile Ala Ala	755								760								765								2304
gcg ggc ggc gat gtg cgc gcc aag ggc aag atc gtg atc gcc acc gtc Ala Gly Gly Asp Val Arg Ala Lys Gly Lys Ile Val Ile Ala Thr Val	770								775								780								2352
aag ggc gac gtg cac gac atc ggc aag aac atc gtg tcg gtg gtc ttg Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser Val Val Leu	790								795								800								2400
cag tgc aat aac ttc gaa gtc gtg aac atg ggc gtg atg gtg ccg tgc Gln Cys Asn Asn Phe Glu Val Val Asn Met Gly Val Met Val Pro Cys	805								810								815								2448
gcc cag atc ctg cag aag gcc aag gac gag aac gcc gac atg atc ggc Ala Gln Ile Leu Gln Lys Ala Lys Asp Glu Asn Ala Asp Met Ile Gly	820								825								830								2496
ctg tcc ggc ctg atc acg ccc agc ctc gaa gag atg gcc tac gtg gct Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala Tyr Val Ala	835								840								845								2544
tca gaa atg cag cgc gac ccc tat ttc cgc gag cgc gcc atg ccg ctg Ser Glu Met Gln Arg Asp Pro Tyr Phe Arg Glu Arg Ala Met Pro Leu	850								855								860								2592
atg ata ggc ggg gcg acc acc agc cgg gtc cat acg gcg gtc aag atc Met Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val Lys Ile	870								875								880								2640
gcg ccc aac tac gac ggt ccg gtg atc tac gtg ccc gat gcc agc cgt Ala Pro Asn Tyr Asp Gly Pro Val Ile Tyr Val Pro Asp Ala Ser Arg	885								890								895								2688
tcg gtc ggc gtg gcg acc agc ctc atg tcc gac cag gcc ccg gcc tat Ser Val Gly Val Ala Thr Ser Leu Met Ser Asp Gln Ala Pro Ala Tyr	900								905								910								2736
ttg gcg gag ctg gcg cag gag tac gag gat gtg cgc cgc tgc cat gcc Leu Ala Glu Leu Ala Gln Glu Tyr Glu Asp Val Arg Arg Cys His Ala	915								920								925								2784
aac cgc aag gcg gtg ccg ctg gtg tcg ctg gcc gag gcg cgc gcg gcg Asn Arg Lys Ala Val Pro Leu Val Ser Leu Ala Glu Ala Arg Ala Ala	930								935								940								2832
cgc ccg cag atc gac tgg tcc ggc tac cag ccg ccg cgc ccc aag ttc Arg Pro Gln Ile Asp Trp Ser Gly Tyr Gln Pro Pro Arg Pro Lys Phe	950								955								960								2880

138

ctg ggc cgg cgc gcc ttc aag agc tac gac ctg gcc gag atc gcg cgc Leu Gly Arg Arg Ala Phe Lys Ser Tyr Asp Leu Ala Glu Ile Ala Arg 965 970 975	2928
tat atc gac tgg ggg ccg ttc ttc cag acg tgg agc ctg ttc ggc ccg Tyr Ile Asp Trp Gly Pro Phe Phe Gln Thr Trp Ser Leu Phe Gly Pro 980 985 990	2976
ttc ccc gcc atc ctg gac gac aag gtg gtg ggc gag cag gcg cgc aag Phe Pro Ala Ile Leu Asp Asp Lys Val Val Gly Glu Gln Ala Arg Lys 995 1000 1005	3024
gtc tac gag gaa ggc cag gcc atg ctc aag cgc atc atc gac ggg cgc Val Tyr Glu Glu Gly Gln Ala Met Leu Lys Arg Ile Ile Asp Gly Arg 1010 1015 1020	3072
tgg ctg acc gcc agc ggc gtg gtc ggc ttc tat ccg gcc aac cgc gtc Trp Leu Thr Ala Ser Gly Val Val Gly Phe Tyr Pro Ala Asn Arg Val 1025 1030 1035 1040	3120
aat gac gaa gac atc gag gtc tac gcg gac gag acg cgc agc gag atg Asn Asp Glu Asp Ile Glu Val Tyr Ala Asp Glu Thr Arg Ser Glu Met 1045 1050 1055	3168
ctg ttc acc tac cgc aac ctg cgc cag cag ggc gtc aag cgc gaa ggc Leu Phe Thr Tyr Arg Asn Leu Arg Gln Gln Gly Val Lys Arg Glu Gly 1060 1065 1070	3216
gtc agc aac aag tgc ctg gcc gac tac atc gcg ccg cgc gac agc ggc Val Ser Asn Lys Cys Leu Ala Asp Tyr Ile Ala Pro Arg Asp Ser Gly 1075 1080 1085	3264
ctg ctc gac tac atc ggc atg ttc gcc gtg acc gcg ggc ctg ggc atc Leu Leu Asp Tyr Ile Gly Met Phe Ala Val Thr Ala Gly Leu Gly Ile 1090 1095 1100	3312
gag aag aaa gag gcc gag ttc cag gcg gcg ctg gac gac tac tcc agc Glu Lys Lys Glu Ala Glu Phe Gln Ala Ala Leu Asp Asp Tyr Ser Ser 1105 1110 1115 1120	3360
atc atg ctg aag tcg ctg gcc gac cgg ctg gcc gag gcg ttc gcc gaa Ile Met Leu Lys Ser Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu 1125 1130 1135	3408
tgc atg cac gcg cgc gtg cgc cgc gac ctg tgg ggc tac gcg gcg gac Cys Met His Ala Arg Val Arg Arg Asp Leu Trp Gly Tyr Ala Ala Asp 1140 1145 1150	3456
gag gcg ctg tcc aac gat gag ctg atc gcc gag aag tac agc ggc atc Glu Ala Leu Ser Asn Asp Glu Leu Ile Ala Glu Lys Tyr Ser Gly Ile 1155 1160 1165	3504
cgg ccg gcg ccc ggc tat ccg gcc tgc ccg gag cac gtg gtc aag acg Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Val Val Lys Thr 1170 1175 1180	3552
gac ctg ttc cgc gtg ctg gac gcc gcc gac gtc gga atg gag ctg acc Asp Leu Phe Arg Val Leu Asp Ala Ala Asp Val Gly Met Glu Leu Thr 1185 1190 1195 1200	3600
gac agc tac gcc atg ttc ccg gcc tcc agc gtc tcg ggg ttc tat ttc	3648



139

Asp Ser Tyr Ala Met Phe Pro Ala Ser Ser Val Ser Gly Phe Tyr Phe  
 1205 1210 1215

agc cac ccc gag tcg cag tat ttc aac gtg ggc aac atc ggc gcc gac 3696  
 Ser His Pro Glu Ser Gln Tyr Phe Asn Val Gly Asn Ile Gly Ala Asp  
 1220 1225 1230

cag ctg gcc gac tac gtg gcg cgc agc ggc cgc gcc gaa gag gac gtg 3744  
 Gln Leu Ala Asp Tyr Val Ala Arg Ser Gly Arg Ala Glu Glu Asp Val  
 1235 1240 1245

cgc cgc acc ctg gcg ccg aac ctg ggc tag 3774  
 Arg Arg Thr Leu Ala Pro Asn Leu Gly  
 1250 1255

&lt;210&gt; 32

&lt;211&gt; 1257

&lt;212&gt; PRT

&lt;213&gt; Bordetella pertussis

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 69 .. 69

&lt;223&gt; All occurrences of Xaa indicate any amino acid

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 93 .. 93

&lt;223&gt; All occurrences of Xaa indicate any amino acid

&lt;400&gt; 32

Val Pro Tyr Pro Arg Ile Pro Phe Pro Leu Ser Ala Tyr Thr His Gly  
 1 5 10 15

Gly Glu Phe Val Arg Gln Leu Asp Lys Arg Ile Leu Ile Leu Asp Gly  
 20 25 30

Ala Met Gly Thr Met Ile Gln Arg Tyr Lys Leu Gly Glu Ala Asp Phe  
 35 40 45

Arg Gly Glu Arg Phe Ala Glu His His Lys Asp Leu Lys Gly Asp Asn  
 50 55 60

Glu Leu Leu Ser Xaa Val Arg Pro Asp Val Ile Ala Glu Ile His Arg  
 65 70 75 80

Gln Tyr Leu Glu Ala Gly Ala Asp Val Ile Glu Thr Xaa Thr Phe Gly  
 85 90 95

Ala Thr Ser Ile Ala Gln Gly Asp Tyr Asp Leu Pro Glu Leu Ala Tyr  
 100 105 110

Glu Met Asn Leu Glu Ser Ala Arg Leu Ala Arg Ala Ala Cys Asp Ala  
 115 120 125

Tyr Ser Thr Pro Glu His Pro Arg Phe Val Ala Gly Ala Leu Gly Pro  
 130 135 140

Gln Pro Lys Thr Ala Ser Ile Ser Pro Asp Val Asn Asp Pro Gly Ala  
 145 150 155 160

140

Arg Asn Val Thr Phe Asp Glu Leu Arg Ala Ala Tyr Val Glu Gln Leu  
 165 170 175  
 Asn Gly Leu Leu Asp Gly Gly Ile Asp Ile Val Leu Ile Glu Thr Ile  
 180 185 190  
 Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Val Glu Glu Ala  
 195 200 205  
 Phe Glu Ala Arg Gly Val Arg Leu Pro Val Met Ile Ser Gly Thr Val  
 210 215 220  
 Thr Asp Ala Ser Gly Arg Ile Leu Ser Gly Gln Thr Val Glu Ala Phe  
 225 230 235 240  
 Trp Asn Ser Val Arg His Ala Arg Pro Val Thr Ile Gly Leu Asn Cys  
 245 250 255  
 Ala Leu Gly Ala Ala Leu Met Arg Pro Tyr Val Ala Glu Leu Ser Lys  
 260 265 270  
 Ile Cys Asp Thr Tyr Val Cys Val Tyr Pro Asn Ala Gly Leu Pro Asn  
 275 280 285  
 Pro Met Ala Glu Thr Gly Phe Asp Glu Thr Pro Ala Asp Thr Ser Ala  
 290 295 300  
 Leu Leu Glu Glu Phe Ala Gln Ala Gly Leu Val Asn Met Ala Gly Gly  
 305 310 315 320  
 Cys Cys Gly Thr Thr Pro Glu His Ile Arg Ala Ile Ala Gly Lys Val  
 325 330 335  
 Ala Ala Leu Thr Pro Arg Ala Val Pro Glu Val Pro Val Lys Thr Arg  
 340 345 350  
 Leu Ser Gly Leu Glu Ala Leu Asn Ile Asp Asp Glu Thr Leu Phe Val  
 355 360 365  
 Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Lys Met Phe Ala Arg  
 370 375 380  
 Leu Val Arg Glu Glu Lys Tyr Asp Glu Ala Leu Ala Val Ala Arg Gln  
 385 390 395 400  
 Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Val Asn Met Asp Glu Ala  
 405 410 415  
 Met Leu Asp Ser Val Ala Cys Met His Arg Phe Leu Asn Leu Ile Ala  
 420 425 430  
 Ser Glu Pro Asp Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys  
 435 440 445  
 Trp Glu Val Ile Glu Thr Gly Leu Lys Cys Val Gln Gly Lys Ala Val  
 450 455 460  
 Val Asn Ser Ile Ser Met Lys Glu Gly Glu Glu Pro Phe Arg His His  
 465 470 475 480  
 Ala Arg Leu Cys Arg Arg Tyr Gly Ala Ala Met Val Val Met Ala Phe  
 485 490 495

141

Asp Glu Gln Gly Gln Ala Asp Ser Leu Glu Arg Arg Lys Glu Ile Cys  
 500 505 510  
 Gly Arg Ala Tyr Arg Ile Leu Val Glu Glu Gly Phe Pro Pro Glu  
 515 520 525  
 Asp Ile Ile Phe Asp Pro Asn Val Phe Ala Val Ala Thr Gly Ile Asp  
 530 535 540  
 Glu His Asn His Tyr Ala Val Asp Phe Ile Glu Gly Ala Arg Trp Ile  
 545 550 555 560  
 Arg Ala Asn Leu Pro His Ala Arg Ile Ser Gly Gly Ile Ser Asn Val  
 565 570 575  
 Ser Phe Ser Phe Arg Gly Asn Glu Pro Met Arg Glu Ala Ile His Thr  
 580 585 590  
 Val Phe Leu Tyr Tyr Ala Ile Glu Ala Gly Leu Thr Met Gly Ile Val  
 595 600 605  
 Asn Ala Gly Gln Leu Gly Val Tyr Ala Asp Leu Ala Pro His Leu Arg  
 610 615 620  
 Asp Leu Val Glu Asp Val Ile Leu Asp Arg Pro Glu Pro Val Gly Arg  
 625 630 635 640  
 Ser Asp Ser Ala Asp Glu Arg Ser Pro Thr Glu Arg Leu Val Gln Phe  
 645 650 655  
 Ala Glu Thr Val Lys Gly Ser Gly Ala Lys Lys Glu Glu Asp Leu Thr  
 660 665 670  
 Trp Arg Thr Gly Ser Val Glu Gln Arg Leu Ala His Ala Leu Val His  
 675 680 685  
 Gly Ile Thr Thr Phe Ile Val Glu Asp Thr Glu Glu Val Arg Gln Gln  
 690 695 700  
 Val Ala Ala Arg Gly Gly Arg Thr Ile Glu Val Ile Glu Gly Pro Leu  
 705 710 715 720  
 Met Asp Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly Lys Met  
 725 730 735  
 Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val  
 740 745 750  
 Ala His Leu Ile Pro Phe Ile Glu Glu Lys Arg Gln Ile Ala Ala  
 755 760 765  
 Ala Gly Gly Asp Val Arg Ala Lys Gly Lys Ile Val Ile Ala Thr Val  
 770 775 780  
 Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser Val Val Leu  
 785 790 795 800  
 Gln Cys Asn Asn Phe Glu Val Val Asn Met Gly Val Met Val Pro Cys  
 805 810 815  
 Ala Gln Ile Leu Gln Lys Ala Lys Asp Glu Asn Ala Asp Met Ile Gly

142

820	825	830
Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala Tyr Val Ala 835	840	845
Ser Glu Met Gln Arg Asp Pro Tyr Phe Arg Glu Arg Ala Met Pro Leu 850	855	860
Met Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val Lys Ile 865	870	875 880
Ala Pro Asn Tyr Asp Gly Pro Val Ile Tyr Val Pro Asp Ala Ser Arg 885	890	895
Ser Val Gly Val Ala Thr Ser Leu Met Ser Asp Gln Ala Pro Ala Tyr 900	905	910
Leu Ala Glu Leu Ala Gln Glu Tyr Glu Asp Val Arg Arg Cys His Ala 915	920	925
Asn Arg Lys Ala Val Pro Leu Val Ser Leu Ala Glu Ala Arg Ala Ala 930	935	940
Arg Pro Gln Ile Asp Trp Ser Gly Tyr Gln Pro Pro Arg Pro Lys Phe 945	950	955 960
Leu Gly Arg Arg Ala Phe Lys Ser Tyr Asp Leu Ala Glu Ile Ala Arg 965	970	975
Tyr Ile Asp Trp Gly Pro Phe Phe Gln Thr Trp Ser Leu Phe Gly Pro 980	985	990
Phe Pro Ala Ile Leu Asp Asp Lys Val Val Gly Glu Gln Ala Arg Lys 995	1000	1005
Val Tyr Glu Glu Gly Gln Ala Met Leu Lys Arg Ile Ile Asp Gly Arg 1010	1015	1020
Trp Leu Thr Ala Ser Gly Val Val Gly Phe Tyr Pro Ala Asn Arg Val 1025	1030	1035 1040
Asn Asp Glu Asp Ile Glu Val Tyr Ala Asp Glu Thr Arg Ser Glu Met 1045	1050	1055
Leu Phe Thr Tyr Arg Asn Leu Arg Gln Gln Gly Val Lys Arg Glu Gly 1060	1065	1070
Val Ser Asn Lys Cys Leu Ala Asp Tyr Ile Ala Pro Arg Asp Ser Gly 1075	1080	1085
Leu Leu Asp Tyr Ile Gly Met Phe Ala Val Thr Ala Gly Leu Gly Ile 1090	1095	1100
Glu Lys Lys Glu Ala Glu Phe Gln Ala Ala Leu Asp Asp Tyr Ser Ser 1105	1110	1115 1120
Ile Met Leu Lys Ser Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu 1125	1130	1135
Cys Met His Ala Arg Val Arg Arg Asp Leu Trp Gly Tyr Ala Ala Asp 1140	1145	1150

143

Glu Ala Leu Ser Asn Asp Glu Leu Ile Ala Glu Lys Tyr Ser Gly Ile  
 1155 1160 1165

Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Val Val Lys Thr  
 1170 1175 1180

Asp Leu Phe Arg Val Leu Asp Ala Ala Asp Val Gly Met Glu Leu Thr  
 1185 1190 1195 1200

Asp Ser Tyr Ala Met Phe Pro Ala Ser Ser Val Ser Gly Phe Tyr Phe  
 1205 1210 1215

Ser His Pro Glu Ser Gln Tyr Phe Asn Val Gly Asn Ile Gly Ala Asp  
 1220 1225 1230

Gln Leu Ala Asp Tyr Val Ala Arg Ser Gly Arg Ala Glu Glu Asp Val  
 1235 1240 1245

Arg Arg Thr Leu Ala Pro Asn Leu Gly  
 1250 1255

&lt;210&gt; 33

&lt;211&gt; 3645

&lt;212&gt; DNA

&lt;213&gt; Chlorobium tepidum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (3642)

&lt;223&gt; RCL00420

&lt;400&gt; 33

gtg ctc gac ggg gcc atg ggc acc atg atc cag agg cat ggc ctc gac 48  
 Val Leu Asp Gly Ala Met Gly Thr Met Ile Gln Arg His Gly Leu Asp  
 1 5 10 15

gaa cag gac tac cgg ggc gag cgt ttc gct tcg cat gac cat ccg ctg 96  
 Glu Gln Asp Tyr Arg Gly Glu Arg Phe Ala Ser His Asp His Pro Leu  
 20 25 30

aag ggc aac aac gac ctt ctt gtc atc acc cgg ccc gac atc atc cgt 144  
 Lys Gly Asn Asn Asp Leu Leu Val Ile Thr Arg Pro Asp Ile Ile Arg  
 35 40 45

tcg atc cac tgc gac ttc ctc gac gcg ggt gcg gac atc atc gag acc 192  
 Ser Ile His Cys Asp Phe Leu Asp Ala Gly Ala Asp Ile Ile Glu Thr  
 50 55 60

tgc acc ttc aac gcc aac ccg atc tcg cag tcg gac tac cag ttg cag 240  
 Cys Thr Phe Asn Ala Asn Pro Ile Ser Gln Ser Asp Tyr Gln Leu Gln  
 65 70 75 80

gac ttg acc cgc gag ctg aac gtg gcg gcg gca aag ata gcc cgc tcg 288  
 Asp Leu Thr Arg Glu Leu Asn Val Ala Ala Lys Ile Ala Arg Ser  
 85 90 95

gca gcg gac gag ttc acc gca aag act ccc gac aag ccg cgt ttc gtg 336  
 Ala Ala Asp Glu Phe Thr Ala Lys Thr Pro Asp Lys Pro Arg Phe Val  
 100 105 110

gcc ggt tcc atc gga ccg acc aac aag acg ctc tcg ctc tcg ccg gac 384

144

Ala Gly Ser Ile Gly Pro Thr Asn Lys Thr Leu Ser Leu Ser Pro Asp	
115 120 125	
gtg aac aac ccc ggc ttc cgc gcc gtc acc ttc cag gag atg gtc gat	432
Val Asn Asn Pro Gly Phe Arg Ala Val Thr Phe Gln Glu Met Val Asp	
130 135 140	
aac tac act gcc cag ctc gaa ggc ttg cac gag gcc ggt gtc gat ctc	480
Asn Tyr Thr Ala Gln Leu Glu Gly Leu His Glu Gly Gly Val Asp Leu	
145 150 155 160	
ttg ctc gtc gag acg gtg ttc gac aca ctg aac tgc aag gcg gcg ctc	528
Leu Leu Val Glu Thr Val Phe Asp Thr Leu Asn Cys Lys Ala Ala Leu	
165 170 175	
tac gct atc gag gag tac gcg gtg aaa acc gcc tgg cag gtg ccc gtg	576
Tyr Ala Ile Glu Glu Tyr Ala Val Lys Thr Gly Trp Gln Val Pro Val	
180 185 190	
atg gtc tcc gcc acg gtg gtg gac gcg agc gcc cgc acc ctc tcc gcc	624
Met Val Ser Gly Thr Val Val Asp Ala Ser Gly Arg Thr Leu Ser Gly	
195 200 205	
caa acc acc gag gcg ttc tgg att tcg att tcg cac atg ccg agt ctg	672
Gln Thr Thr Glu Ala Phe Trp Ile Ser Ile Ser His Met Pro Ser Leu	
210 215 220	
ctc tcg gtc gcc ctg aac tgc gca ctc gcc tcc aag cag atg cgc ccc	720
Leu Ser Val Gly Leu Asn Cys Ala Leu Gly Ser Lys Gln Met Arg Pro	
225 230 235 240	
ttc atc gag gcg ctc tcg aac atc gcc gaa agc tac gtc agc gtc tat	768
Phe Ile Glu Ala Leu Ser Asn Ile Ala Glu Ser Tyr Val Ser Val Tyr	
245 250 255	
ccc aac gcg gcc ctg ccg aat gag ttc gcc gag tac gac gac tcc ccc	816
Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp Asp Ser Pro	
260 265 270	
gag tac atg gcc gcg cag atc gcg gcc ttc gcc gaa tca gcc ttc gtg	864
Glu Tyr Met Ala Ala Gln Ile Ala Gly Phe Ala Glu Ser Gly Phe Val	
275 280 285	
aac atc gtc gcc gcc tgc tgc gcc acc acg ccg acg cac atc cgc gcc	912
Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Thr His Ile Arg Ala	
290 295 300	
att gcc gaa gcg gtc aag act ctc ccg ccg aga aag cgc ccc gcc aac	960
Ile Ala Glu Ala Val Lys Thr Leu Pro Pro Arg Lys Arg Pro Ala Asn	
305 310 315 320	
aag cac gtg ctg agg ctc tcc gcc ctc gaa ccg ctc gtg gtt gac gaa	1008
Lys His Val Leu Arg Leu Ser Gly Leu Glu Pro Leu Val Val Asp Glu	
325 330 335	
acc acc gcc ttc atc aac gtc gcc gag cgc acc aac gtc acc ggt tcg	1056
Thr Thr Gly Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser	
340 345 350	
cgc aag ttc gcc cgc ctc atc aag gag gcc aat tac gac gaa gcg ctc	1104
Arg Lys Phe Ala Arg Leu Ile Lys Glu Ala Asn Tyr Asp Glu Ala Leu	
355 360 365	

## 145

tcc att gcc cgc cag cag gtc gag aac ggc gcg cag gtg atc gac gtg Ser Ile Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile Asp Val 370 375 380	1152
aac ctc gac gaa gga atg ctc gac tcc gaa aag gtg atc gtc gaa ttc Asn Leu Asp Glu Gly Met Leu Asp Ser Glu Lys Val Ile Val Glu Phe 385 390 395 400	1200
ctg aac ctc atc gcc tcc gag cct gag atc gcc aag gtg ccg gtg atg Leu Asn Leu Ile Ala Ser Glu Pro Glu Ile Ala Lys Val Pro Val Met 405 410 415	1248
atc gac tcg tcg aaa tgg tcg gtc atc gaa aac ggc ctg cgc tgc acc Ile Asp Ser Ser Lys Trp Ser Val Ile Glu Asn Gly Leu Arg Cys Thr 420 425 430	1296
cag ggc aag agc atc gtc aac tcg atc agc ctc aag gag ggc gag gag Gln Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu 435 440 445	1344
ctg ttc aag gag cgc gct cgc aag atc atg caa tac ggc gcg gcg gcg Leu Phe Lys Glu Arg Ala Arg Lys Ile Met Gln Tyr Gly Ala Ala Ala 450 455 460	1392
gtg gtc atg gcc ttc gac gag cag ggc cag gcc gac agc ctg cac cgc Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Ser Leu His Arg 465 470 475 480	1440
cgc atc gag att tgc agc cgc gcc tac aaa att ctc acc gaa gag gtg Arg Ile Glu Ile Cys Ser Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val 485 490 495	1488
ggc ttc ccg ccg gag gac atc atc ttt gac ccg aac gtg ctg acc gtg Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val 500 505 510	1536
gcc acc ggc atc gac gag cac aac aac tac gcg ctc gac ttc atc gaa Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile Glu 515 520 525	1584
agc gtg cgc tgg atc aag cag aac ctg ccg cac gcg aag gtc tcc ggc Ser Val Arg Trp Ile Lys Gln Asn Leu Pro His Ala Lys Val Ser Gly 530 535 540	1632
ggc atc agc aac gtt tcg ttc tcc ttc cgc ggc aac gag ccg gtg cgc Gly Ile Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg 545 550 555 560	1680
gag gcg atg cac acc gcg ttc ctc tac cac gcc atc cac gcc ggt ctc Glu Ala Met His Thr Ala Phe Leu Tyr His Ala Ile His Ala Gly Leu 565 570 575	1728
gac atg ggc atc gtc aac gcc gcc cag ctt ggc atc tac gaa gag atc Asp Met Gly Ile Val Asn Ala Ala Gln Leu Gly Ile Tyr Glu Glu Ile 580 585 590	1776
gac ccg gag ctt ctt gtc tat gtc gag gac gtg ctg ctg aac cgc cgc Asp Pro Glu Leu Leu Val Tyr Val Glu Asp Val Leu Leu Asn Arg Arg 595 600 605	1824
gac gac gcc acc gag cgg ctc gtg gcg ttc gct gaa acg atc cgc gac	1872

146

Asp Asp Ala Thr Glu Arg Leu Val Ala Phe Ala Glu Thr Ile Arg Asp 610 615 620	
ggc ggc gaa aag gcc gag gcc aag aac gcc gaa tgg cgc aac gcc ccg Gly Gly Glu Lys Ala Glu Ala Lys Asn Ala Glu Trp Arg Asn Ala Pro 625 630 635 640	1920
gtc gag gag cgg ctg aaa cac gcg ctc gtc aag ggc atc gtt gac tac Val Glu Glu Arg Leu Lys His Ala Leu Val Lys Gly Ile Val Asp Tyr 645 650 655	1968
atc gac gag gac acc gaa gag gcc cgc cag ctc tac ccg agt ccg ctg Ile Asp Glu Asp Thr Glu Glu Ala Arg Gln Leu Tyr Pro Ser Pro Leu 660 665 670	2016
gag gtg atc gag ggg ccg ctc atg aac ggc atg aac cac gtc ggc gac Glu Val Ile Glu Gly Pro Leu Met Asn Gly Met Asn His Val Gly Asp 675 680 685	2064
ctc ttc gcc gaa ggc aag atg ttc ctg cca cag gtg gtc aaa agc gcc Leu Phe Ala Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala 690 695 700	2112
cgc gtc atg aag cgc tcg gta gct gcg ctg att ccc tat atc gag gag Arg Val Met Lys Arg Ser Val Ala Ala Leu Ile Pro Tyr Ile Glu Glu 705 710 715 720	2160
gag aag tcg aaa aac tgc gac acg agc gcc aaa gcc aag gtg ctg ctc Glu Lys Ser Lys Asn Cys Asp Thr Ser Ala Lys Ala Lys Val Leu Leu 725 730 735	2208
gcc acg gtg aag ggc gac gtg cac gac atc ggc aag aac atc gtg tcg Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser 740 745 750	2256
gtg gtg ctt gcc tgc aac aac ttc gac gtg atc gac atc ggc gtc atg Val Val Leu Ala Cys Asn Asn Phe Asp Val Ile Asp Ile Gly Val Met 755 760 765	2304
atg cca tgc gac aag att ctc gaa gcg ctg gca gaa cac aag ccc gac Met Pro Cys Asp Lys Ile Leu Glu Ala Leu Ala Glu His Lys Pro Asp 770 775 780	2352
gtg ctc ggc ctc tcc ggc ctc atc acc ccg tcg ctc gaa gag atg gcg Val Leu Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala 785 790 795 800	2400
cac gtg gcc aaa gag atg gag cgg ctc ggc atg aac att ccg ctc atc His Val Ala Lys Glu Met Glu Arg Leu Gly Met Asn Ile Pro Leu Ile 805 810 815	2448
atc ggc ggc gcg acc acc tcg aag gtg cac acg gcg gtg aaa ctc gcg Ile Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Leu Ala 820 825 830	2496
ccc tgc tac ccc agc ggc gcg gta gta cac gtg ctc gac gcc tcg cgc Pro Cys Tyr Pro Ser Gly Ala Val Val His Val Leu Asp Ala Ser Arg 835 840 845	2544
agc gtg ccg gtg gtc agc aac ctc tgc aac ccc gcc cag cgc gac agc Ser Val Pro Val Val Ser Asn Leu Cys Asn Pro Ala Gln Arg Asp Ser 850 855 860	2592



tat atc gcg gcg ctg aag gat gag cag gag gcg atg cgc aag agc cac Tyr Ile Ala Ala Leu Lys Asp Glu Gln Glu Ala Met Arg Lys Ser His 865 870 875 880	2640
gcc gag cgc atg gcg gca aaa aag tac gtc tcg ctc gac gcc gcc cgc Ala Glu Arg Met Ala Ala Lys Lys Tyr Val Ser Leu Asp Ala Ala Arg 885 890 895	2688
gac aac cgc ctc acc att gac tgg gag gcc gaa acc atc gac aag ccc Asp Asn Arg Leu Thr Ile Asp Trp Glu Ala Glu Thr Ile Asp Lys Pro 900 905 910	2736
gcc cag act ggc gtc acc gtg ctg gag gat gtc acc gtc ggc gcg ctc Ala Gln Thr Gly Val Thr Val Leu Glu Asp Val Thr Val Gly Ala Leu 915 920 925	2784
cgc ccg tat atc gac tgg gca mcc ttc ttc tgg agc tgg gag ctg cac Arg Pro Tyr Ile Asp Trp Ala Xaa Phe Phe Trp Ser Trp Glu Leu His 930 935 940	2832
ggc gtc tat ccg cag att ctg gag gat gaa aag gtc ggc gag gag gca Gly Val Tyr Pro Gln Ile Leu Glu Asp Glu Lys Val Gly Glu Glu Ala 945 950 955 960	2880
acc aaa ctc ttc aac gac gcc acc gct ctg ctc gac cgg atc gac agc Thr Lys Leu Phe Asn Asp Ala Thr Ala Leu Leu Asp Arg Ile Asp Ser 965 970 975	2928
gaa aag ctg ctc ggc atc aaa ggc gtg gcg ggc atc ttc ccg gcc aac Glu Lys Leu Leu Gly Ile Lys Gly Val Ala Gly Ile Phe Pro Ala Asn 980 985 990	2976
agc atc ggc gac gac atc ttc gtc tat gcg gat gac gag cgc tcg ata Ser Ile Gly Asp Asp Ile Phe Val Tyr Ala Asp Asp Glu Arg Ser Ile 995 1000 1005	3024
atc cgc acc gtg ctg cac acc ctg cgc cag caa ggc gaa aag cac ggc Ile Arg Thr Val Leu His Thr Leu Arg Gln Gln Gly Glu Lys His Gly 1010 1015 1020	3072
gaa gcg aac ctc gcg ctg gcg gac ttc gtg gcc ccg cgc gaa agc ggc Glu Ala Asn Leu Ala Leu Ala Asp Phe Val Ala Pro Arg Glu Ser Gly 1025 1030 1035 1040	3120
gtc aac gac tgg atc ggc tgc ttc acc gta acc gcc gga ctc ggc atc Val Asn Asp Trp Ile Gly Cys Phe Thr Val Thr Ala Gly Leu Gly Ile 1045 1050 1055	3168
cag aat ttg ctc gac gag ttc aca gca gag aac gac gac tac cac cgc Gln Asn Leu Leu Asp Glu Phe Thr Ala Glu Asn Asp Asp Tyr His Arg 1060 1065 1070	3216
atc atg aca cag gcg ctc gcc gac cga ctg gcc gaa gcg ttc gca gag Ile Met Thr Gln Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu 1075 1080 1085	3264
atg ctg cac gaa aag gtg cgc cgc gaa ctc tgg ggc tac gcg ccc ggc Met Leu His Glu Lys Val Arg Arg Glu Leu Trp Gly Tyr Ala Pro Gly 1090 1095 1100	3312
gaa atc ctc ggc aac gaa gag ctg atc gcc gaa aag tac cga ggc atc	3360

148

Glu Ile Leu Gly Asn Glu Glu Leu Ile Ala Glu Lys Tyr Arg Gly Ile  
 1105 1110 1115 1120  
 cgc ccc gcc ccc ggc tac ccc gcc tgc ccg gat cac acc gaa aag gca 3408  
 Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys Ala  
 1125 1130 1135  
 atc atc ttc gac ctg ctc aac gct gaa gcg gcc acc ggc gtc acg ctg 3456  
 Ile Ile Phe Asp Leu Leu Asn Ala Glu Ala Ala Thr Gly Val Thr Leu  
 1140 1145 1150  
 acg gaa act ttc gcg atg aac ccc gca gcc tca gtc tgc ggc ctc tac 3504  
 Thr Glu Thr Phe Ala Met Asn Pro Ala Ala Ser Val Cys Gly Leu Tyr  
 1155 1160 1165  
 ttc gcc aac ccg gcc tcg aaa tac ttc gta ctc ggc aag att ggt aag 3552  
 Phe Ala Asn Pro Ala Ser Lys Tyr Phe Val Leu Gly Lys Ile Gly Lys  
 1170 1175 1180  
 gat cag gtc gaa gac tac gcc aac cgc aaa ggg ctg gaa gta gca gaa 3600  
 Asp Gln Val Glu Asp Tyr Ala Asn Arg Lys Gly Leu Glu Val Ala Glu  
 1185 1190 1195 1200  
 gcc gag aag tgg ctc gcg ccc tcg ctg aac tac gat cca gcg 3642  
 Ala Glu Lys Trp Leu Ala Pro Ser Leu Asn Tyr Asp Pro Ala  
 1205 1210  
 taa 3645  
 <210> 34  
 <211> 1214  
 <212> PRT  
 <213> Chlorobium tepidum  
 <220>  
 <221> unsure  
 <222> 936 .. 936  
 <223> All occurrences of Xaa indicate any amino acid  
 <400> 34  
 Val Leu Asp Gly Ala Met Gly Thr Met Ile Gln Arg His Gly Leu Asp  
 1 5 10 15  
 Glu Gln Asp Tyr Arg Gly Glu Arg Phe Ala Ser His Asp His Pro Leu  
 20 25 30  
 Lys Gly Asn Asn Asp Leu Leu Val Ile Thr Arg Pro Asp Ile Ile Arg  
 35 40 45  
 Ser Ile His Cys Asp Phe Leu Asp Ala Gly Ala Asp Ile Ile Glu Thr  
 50 55 60  
 Cys Thr Phe Asn Ala Asn Pro Ile Ser Gln Ser Asp Tyr Gln Leu Gln  
 65 70 75 80  
 Asp Leu Thr Arg Glu Leu Asn Val Ala Ala Ala Lys Ile Ala Arg Ser  
 85 90 95  
 Ala Ala Asp Glu Phe Thr Ala Lys Thr Pro Asp Lys Pro Arg Phe Val  
 100 105 110

149

Ala Gly Ser Ile Gly Pro Thr Asn Lys Thr Leu Ser Leu Ser Pro Asp  
 115 120 125  
 Val Asn Asn Pro Gly Phe Arg Ala Val Thr Phe Gln Glu Met Val Asp  
 130 135 140  
 Asn Tyr Thr Ala Gln Leu Glu Gly Leu His Glu Gly Gly Val Asp Leu  
 145 150 155 160  
 Leu Leu Val Glu Thr Val Phe Asp Thr Leu Asn Cys Lys Ala Ala Leu  
 165 170 175  
 Tyr Ala Ile Glu Glu Tyr Ala Val Lys Thr Gly Trp Gln Val Pro Val  
 180 185 190  
 Met Val Ser Gly Thr Val Val Asp Ala Ser Gly Arg Thr Leu Ser Gly  
 195 200 205  
 Gln Thr Thr Glu Ala Phe Trp Ile Ser Ile Ser His Met Pro Ser Leu  
 210 215 220  
 Leu Ser Val Gly Leu Asn Cys Ala Leu Gly Ser Lys Gln Met Arg Pro  
 225 230 235 240  
 Phe Ile Glu Ala Leu Ser Asn Ile Ala Glu Ser Tyr Val Ser Val Tyr  
 245 250 255  
 Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp Asp Ser Pro  
 260 265 270  
 Glu Tyr Met Ala Ala Gln Ile Ala Gly Phe Ala Glu Ser Gly Phe Val  
 275 280 285  
 Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Thr His Ile Arg Ala  
 290 295 300  
 Ile Ala Glu Ala Val Lys Thr Leu Pro Pro Arg Lys Arg Pro Ala Asn  
 305 310 315 320  
 Lys His Val Leu Arg Leu Ser Gly Leu Glu Pro Leu Val Val Asp Glu  
 325 330 335  
 Thr Thr Gly Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser  
 340 345 350  
 Arg Lys Phe Ala Arg Leu Ile Lys Glu Ala Asn Tyr Asp Glu Ala Leu  
 355 360 365  
 Ser Ile Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile Asp Val  
 370 375 380  
 Asn Leu Asp Glu Gly Met Leu Asp Ser Glu Lys Val Ile Val Glu Phe  
 385 390 395 400  
 Leu Asn Leu Ile Ala Ser Glu Pro Glu Ile Ala Lys Val Pro Val Met  
 405 410 415  
 Ile Asp Ser Ser Lys Trp Ser Val Ile Glu Asn Gly Leu Arg Cys Thr  
 420 425 430  
 Gln Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu  
 435 440 445

## 150

Leu Phe Lys Glu Arg Ala Arg Lys Ile Met Gln Tyr Gly Ala Ala Ala  
 450 455 460  
 Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Ser Leu His Arg  
 465 470 475 480  
 Arg Ile Glu Ile Cys Ser Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val  
 485 490 495  
 Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val  
 500 505 510  
 Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile Glu  
 515 520 525  
 Ser Val Arg Trp Ile Lys Gln Asn Leu Pro His Ala Lys Val Ser Gly  
 530 535 540  
 Gly Ile Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg  
 545 550 555 560  
 Glu Ala Met His Thr Ala Phe Leu Tyr His Ala Ile His Ala Gly Leu  
 565 570 575  
 Asp Met Gly Ile Val Asn Ala Ala Gln Leu Gly Ile Tyr Glu Glu Ile  
 580 585 590  
 Asp Pro Glu Leu Leu Val Tyr Val Glu Asp Val Leu Leu Asn Arg Arg  
 595 600 605  
 Asp Asp Ala Thr Glu Arg Leu Val Ala Phe Ala Glu Thr Ile Arg Asp  
 610 615 620  
 Gly Gly Glu Lys Ala Glu Ala Lys Asn Ala Glu Trp Arg Asn Ala Pro  
 625 630 635 640  
 Val Glu Glu Arg Leu Lys His Ala Leu Val Lys Gly Ile Val Asp Tyr  
 645 650 655  
 Ile Asp Glu Asp Thr Glu Glu Ala Arg Gln Leu Tyr Pro Ser Pro Leu  
 660 665 670  
 Glu Val Ile Glu Gly Pro Leu Met Asn Gly Met Asn His Val Gly Asp  
 675 680 685  
 Leu Phe Ala Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala  
 690 695 700  
 Arg Val Met Lys Arg Ser Val Ala Ala Leu Ile Pro Tyr Ile Glu Glu  
 705 710 715 720  
 Glu Lys Ser Lys Asn Cys Asp Thr Ser Ala Lys Ala Lys Val Leu Leu  
 725 730 735  
 Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser  
 740 745 750  
 Val Val Leu Ala Cys Asn Asn Phe Asp Val Ile Asp Ile Gly Val Met  
 755 760 765  
 Met Pro Cys Asp Lys Ile Leu Glu Ala Leu Ala Glu His Lys Pro Asp

151

770		775		780
Val Leu Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala				
785		790		800
His Val Ala Lys Glu Met Glu Arg Leu Gly Met Asn Ile Pro Leu Ile				
	805		810	815
Ile Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Leu Ala				
	820		825	830
Pro Cys Tyr Pro Ser Gly Ala Val Val His Val Leu Asp Ala Ser Arg				
	835		840	845
Ser Val Pro Val Val Ser Asn Leu Cys Asn Pro Ala Gln Arg Asp Ser				
	850		855	860
Tyr Ile Ala Ala Leu Lys Asp Glu Gln Glu Ala Met Arg Lys Ser His				
	865		870	875
				880
Ala Glu Arg Met Ala Ala Lys Lys Tyr Val Ser Leu Asp Ala Ala Arg				
	885		890	895
Asp Asn Arg Leu Thr Ile Asp Trp Glu Ala Glu Thr Ile Asp Lys Pro				
	900		905	910
Ala Gln Thr Gly Val Thr Val Leu Glu Asp Val Thr Val Gly Ala Leu				
	915		920	925
Arg Pro Tyr Ile Asp Trp Ala Xaa Phe Phe Trp Ser Trp Glu Leu His				
	930		935	940
Gly Val Tyr Pro Gln Ile Leu Glu Asp Glu Lys Val Gly Glu Glu Ala				
	945		950	955
				960
Thr Lys Leu Phe Asn Asp Ala Thr Ala Leu Leu Asp Arg Ile Asp Ser				
	965		970	975
Glu Lys Leu Leu Gly Ile Lys Gly Val Ala Gly Ile Phe Pro Ala Asn				
	980		985	990
Ser Ile Gly Asp Asp Ile Phe Val Tyr Ala Asp Asp Glu Arg Ser Ile				
	995		1000	1005
Ile Arg Thr Val Leu His Thr Leu Arg Gln Gln Gly Glu Lys His Gly				
	1010		1015	1020
Glu Ala Asn Leu Ala Leu Ala Asp Phe Val Ala Pro Arg Glu Ser Gly				
	1025		1030	1035
				1040
Val Asn Asp Trp Ile Gly Cys Phe Thr Val Thr Ala Gly Leu Gly Ile				
	1045		1050	1055
Gln Asn Leu Leu Asp Glu Phe Thr Ala Glu Asn Asp Asp Tyr His Arg				
	1060		1065	1070
Ile Met Thr Gln Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu				
	1075		1080	1085
Met Leu His Glu Lys Val Arg Arg Glu Leu Trp Gly Tyr Ala Pro Gly				
	1090		1095	1100

152

Glu Ile Leu Gly Asn Glu Glu Leu Ile Ala Glu Lys Tyr Arg Gly Ile  
 1105 1110 1115 1120

Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys Ala  
 1125 1130 1135

Ile Ile Phe Asp Leu Leu Asn Ala Glu Ala Ala Thr Gly Val Thr Leu  
 1140 1145 1150

Thr Glu Thr Phe Ala Met Asn Pro Ala Ala Ser Val Cys Gly Leu Tyr  
 1155 1160 1165

Phe Ala Asn Pro Ala Ser Lys Tyr Phe Val Leu Gly Lys Ile Gly Lys  
 1170 1175 1180

Asp Gln Val Glu Asp Tyr Ala Asn Arg Lys Gly Leu Glu Val Ala Glu  
 1185 1190 1195 1200

Ala Glu Lys Trp Leu Ala Pro Ser Leu Asn Tyr Asp Pro Ala  
 1205 1210

&lt;210&gt; 35

&lt;211&gt; 3777

&lt;212&gt; DNA

<213> *Deinococcus radiodurans*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(3774)

&lt;223&gt; RDR02645

&lt;400&gt; 35

atg agc cat cac cca gaa gcg tcg gct tcc gcc aat ccg tcc atc aac 48  
 Met Ser His His Pro Glu Ala Ser Ala Ser Ala Asn Pro Ser Ile Asn  
 1 5 10 15

cat caa ccg tcc acc atc acc gag gcc gcc cgc cag cgc atc ctg att 96  
 His Gln Pro Ser Thr Ile Thr Glu Ala Ala Arg Gln Arg Ile Leu Ile  
 20 25 30

ctc gac ggc gcc tgg ggt acg cag ctt cag cga gcc aac ctc acc gaa 144  
 Leu Asp Gly Ala Trp Gly Thr Gln Leu Gln Arg Ala Asn Leu Thr Glu  
 35 40 45

gcg gac ttc cgc tgg gac gaa gcc gac ccc acg cgg atg tac cgg ggc 192  
 Ala Asp Phe Arg Trp Asp Glu Ala Asp Pro Thr Arg Met Tyr Arg Gly  
 50 55 60

aac ttc gac ctg ctg caa ctg acc aag cct gac gtg att cgc gcc gtg 240  
 Asn Phe Asp Leu Leu Gln Leu Thr Lys Pro Asp Val Ile Arg Ala Val  
 65 70 75 80

cac cgc gcc tat ttc gag gcc gga gcg gac atc gcc agc acc aat acc 288  
 His Arg Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ala Ser Thr Asn Thr  
 85 90 95

ttc aac tcc acg acc atc tcg cag gcg gat tac ggc acc gag gca ctg 336  
 Phe Asn Ser Thr Thr Ile Ser Gln Ala Asp Tyr Gly Thr Glu Ala Leu  
 100 105 110

gcc tac gcc atg aac cgc gag ggg gca agg ctg gcc cgc gaa gtc gcc 384

153

Ala Tyr Ala Met Asn Arg Glu Gly Ala Arg Leu Ala Arg Glu Val Ala	
115 120 125	
gac gag ttc gag gcg cgc gac ggc aaa aag cgc tgg gtg gcg ggg agt	432
Asp Glu Phe Glu Ala Arg Asp Gly Lys Lys Arg Trp Val Ala Gly Ser	
130 135 140	
gtc ggt ccc acc aac cgc acc gcg acc ctt tct ccc gac gtg gag cgg	480
Val Gly Pro Thr Asn Arg Thr Ala Thr Leu Ser Pro Asp Val Glu Arg	
145 150 155 160	
ccc gag ttc cgc aac gtg acc tac gac gac ctc gtg gcg gcg tac tcg	528
Pro Glu Phe Arg Asn Val Thr Tyr Asp Asp Leu Val Ala Ala Tyr Ser	
165 170 175	
gag gcc atc acc ggg ttg atg gaa ggt ggc gcg gac ctg ctg ctc att	576
Glu Ala Ile Thr Gly Leu Met Glu Gly Gly Ala Asp Leu Leu Leu Ile	
180 185 190	
gaa acg gtg ttt gac acg ctg aac gcc aaa gcc gcg ctg ttt gcc gcg	624
Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Leu Phe Ala Ala	
195 200 205	
cag gac gtg ttc gcg gcg cag ggg cgc gag ctg ccg gtc atg ctc tcg	672
Gln Asp Val Phe Ala Ala Gln Gly Arg Glu Leu Pro Val Met Leu Ser	
210 215 220	
ggc acc atc acc gac gcc tcg ggc cgc acg ctg agc ggg cag acg ccc	720
Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Pro	
225 230 235 240	
gaa gcc ttc gcg gtg agc acc gag cac gcc ggc ctc ttt tcg ctg ggc	768
Glu Ala Phe Ala Val Ser Thr Glu His Ala Gly Leu Phe Ser Leu Gly	
245 250 255	
ctg aac tgc gcg ctg ggc gcc gac ctg ctg cgg ccc cac ctg cgc gca	816
Leu Asn Cys Ala Leu Gly Ala Asp Leu Leu Arg Pro His Leu Arg Ala	
260 265 270	
att gcg gcg aac acg gag gcg ctg gtg tcg gtt cac ccc aac gcg ggc	864
Ile Ala Ala Asn Thr Glu Ala Leu Val Ser Val His Pro Asn Ala Gly	
275 280 285	
ctc ccc aac gcc ttc ggg gaa tac gac gaa acg ccc gaa cac acg gcg	912
Leu Pro Asn Ala Phe Gly Glu Tyr Asp Glu Thr Pro Glu His Thr Ala	
290 295 300	
gcg gtg ctg gcc gac ttc gcc cgc gag ggg ctg gtc aac atc gtg ggc	960
Ala Val Leu Ala Asp Phe Ala Arg Glu Gly Leu Val Asn Ile Val Gly	
305 310 315 320	
ggc tgc tgc ggc acc aca ccc gag cac atc aaa gcg att gcg gag gcg	1008
Gly Cys Cys Gly Thr Thr Pro Glu His Ile Lys Ala Ile Ala Glu Ala	
325 330 335	
gtg aag gac att ccc ccg cgc cag gcg ctg caa ctg ccg cct tac ctg	1056
Val Lys Asp Ile Pro Pro Arg Gln Ala Leu Gln Leu Pro Pro Tyr Leu	
340 345 350	
cgc ctc agc ggc ctc gaa gcc ttc acc ctg acg ccg gaa acc aac ttc	1104
Arg Leu Ser Gly Leu Glu Ala Phe Thr Leu Thr Pro Glu Thr Asn Phe	
355 360 365	

154

gtc aac gtg ggc gag cgc acc aac gtg acc ggc agt ccc aag ttc agc Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Pro Lys Phe Ser 370 375 380	1152
aag gcg att ctg gcc ggc gac tac gac gca ggg ctc aag att gcc cgc Lys Ala Ile Leu Ala Gly Asp Tyr Asp Ala Gly Leu Lys Ile Ala Arg 385 390 395 400	1200
cag cag gtg acg aac ggc gcg caa atc gtg gac atc aac ttc gac gag Gln Gln Val Thr Asn Gly Ala Gln Ile Val Asp Ile Asn Phe Asp Glu 405 410 415	1248
ggg atg ctc gac ggc gaa gga gcg atg gtc aag ttc ctc aac ctg ctc Gly Met Leu Asp Gly Glu Gly Ala Met Val Lys Phe Leu Asn Leu Leu 420 425 430	1296
gcc ggg gag ccg gac atc tcg cgc gtg ccc ctg atg ctc gac tcg tcc Ala Gly Glu Pro Asp Ile Ser Arg Val Pro Leu Met Leu Asp Ser Ser 435 440 445	1344
aag tgg gag att ctg gaa gcg ggg ctg cgg cgg gtg cag ggc aag gca Lys Trp Glu Ile Leu Glu Ala Gly Leu Arg Arg Val Gln Gly Lys Ala 450 455 460	1392
gtc gtc aac tcc atc tcg ctc aag gac ggc gag gcc agg ttt ctg gaa Val Val Asn Ser Ile Ser Leu Lys Asp Gly Glu Ala Arg Phe Leu Glu 465 470 475 480	1440
cgc gcc cgg ctg ctg cgg cgc tac ggg gcg gcg gcg gtg gtc atg gcc Arg Ala Arg Leu Leu Arg Arg Tyr Gly Ala Ala Ala Val Val Met Ala 485 490 495	1488
ttc gac gaa cag gga cag gcc gac aac ctc gcc cga cgc cgg gag att Phe Asp Glu Gln Gly Gln Ala Asp Asn Leu Ala Arg Arg Arg Glu Ile 500 505 510	1536
ctg ggc cgc gcg tat agg ctg ctg acc gag cag gcg gac ttt ccg ccg Leu Gly Arg Ala Tyr Arg Leu Leu Thr Glu Gln Ala Asp Phe Pro Pro 515 520 525	1584
cag gac atc att ttc gac ccc aac gtg ctg acc gtt gcc acc ggc atc Gln Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val Ala Thr Gly Ile 530 535 540	1632
gag gaa cac gac cgc tac gcg ctg gac ttt atc gag gcg acg cgc tgg Glu Glu His Asp Arg Tyr Ala Leu Asp Phe Ile Glu Ala Thr Arg Trp 545 550 555 560	1680
att aaa gaa aac ctg ccg gcg gcg aag gtg tcg ggc ggg att tcc aac Ile Lys Glu Asn Leu Pro Ala Ala Lys Val Ser Gly Gly Ile Ser Asn 565 570 575	1728
gtc tcg ttc agc ttc cgg ggc aac aac cac gtg cgc gag gcg atg cac Val Ser Phe Ser Phe Arg Gly Asn Asn His Val Arg Glu Ala Met His 580 585 590	1776
gcg gtg ttt ctg tac cac gcc atc cgc gcc ggg ctg gac atg ggc atc Ala Val Phe Leu Tyr His Ala Ile Arg Ala Gly Leu Asp Met Gly Ile 595 600 605	1824
gtg aac gcg ggg atg ctg gcg gtg tac gag gac atc gag ccg gag ctg	1872



Val	Asn	Ala	Gly	Met	Leu	Ala	Val	Tyr	Glu	Asp	Ile	Glu	Pro	Glu	Leu	
610						615						620				
cgc	gag	gcc	gtc	gag	gac	gtc	att	ctg	gct	cgc	cg	ccg	gac	gcc	acc	1920
Arg	Glu	Ala	Val	Glu	Asp	Val	Ile	Leu	Ala	Arg	Arg	Pro	Asp	Ala	Thr	
625					630					635					640	
gag	cg	ttg	ctg	acg	ctg	gcc	gac	cgc	tac	aag	gac	atc	aag	cgc	gaa	1968
Glu	Arg	Leu	Leu	Thr	Leu	Ala	Asp	Arg	Tyr	Lys	Asp	Ile	Lys	Arg	Glu	(
				645					650					655		
agt	gcc	gcc	cag	agc	gcc	tgg	cgc	gac	ctg	ccg	gtg	cag	gaa	cgg	ctg	2016
Ser	Ala	Ala	Gln	Ser	Ala	Trp	Arg	Asp	Leu	Pro	Val	Gln	Glu	Arg	Leu	
			660					665					670			
cgg	cac	gca	ctg	gtg	cag	ggc	gtc	gcc	gac	cac	gtg	gat	gag	gac	gcc	2064
Arg	His	Ala	Leu	Val	Gln	Gly	Val	Ala	Asp	His	Val	Asp	Glu	Asp	Ala	
		675					680					685				
gag	gcc	gcc	tat	cag	gaa	ctc	ggc	agc	ccg	ctg	gcc	gtc	atc	gaa	ggc	2112
Glu	Ala	Ala	Tyr	Gln	Glu	Leu	Gly	Ser	Pro	Leu	Ala	Val	Ile	Glu	Gly	
		690				695					700					
ccg	ctg	atg	gac	ggc	atg	aac	gtg	gtg	ggc	gac	ctc	ttc	ggc	gcg	ggg	2160
Pro	Leu	Met	Asp	Gly	Met	Asn	Val	Val	Gly	Asp	Leu	Phe	Gly	Ala	Gly	
705					710					715					720	
aaa	atg	ttc	ctg	ccg	cag	gtg	gtc	aaa	tcc	gcc	cgc	gtg	atg	aaa	aag	2208
Lys	Met	Phe	Leu	Pro	Gln	Val	Val	Lys	Ser	Ala	Arg	Val	Met	Lys	Lys	
				725					730					735		
gca	gtg	gcc	tac	ctc	acg	ccc	tat	ctg	gaa	gcg	gag	aag	gcg	gaa	agc	2256
Ala	Val	Ala	Tyr	Leu	Thr	Pro	Tyr	Leu	Glu	Ala	Glu	Lys	Ala	Glu	Ser	
			740					745					750			
tcc	agc	aag	ggc	aag	gta	ctg	ctg	gcg	acc	gtc	aag	ggc	gat	gtg	cac	2304
Ser	Ser	Lys	Gly	Lys	Val	Leu	Leu	Ala	Thr	Val	Lys	Gly	Asp	Val	His	
		755				760						765				
gac	atc	ggc	aag	aac	atc	gtg	ggc	gtg	gtg	ctc	gcc	tgc	aac	ggc	tat	2352
Asp	Ile	Gly	Lys	Asn	Ile	Val	Gly	Val	Val	Leu	Ala	Cys	Asn	Gly	Tyr	
		770				775					780					
cag	gtg	acc	gac	ctc	ggc	gtg	atg	gtg	ccg	ggc	gag	aag	att	ctg	gac	2400
Gln	Val	Thr	Asp	Leu	Gly	Val	Met	Val	Pro	Gly	Glu	Lys	Ile	Leu	Asp	
785					790					795					800	
gaa	gcc	gag	cgg	ctc	ggc	gtg	atg	gtg	atc	ggc	ctg	agc	ggg	ctg	att	2448
Glu	Ala	Glu	Arg	Leu	Gly	Ala	Asp	Val	Ile	Gly	Leu	Ser	Gly	Leu	Ile	
				805					810							

156

cac gtg ctg gac gcc agc cgc gcc gtg acc gtg acc aac gac ctg ctg His Val Leu Asp Ala Ser Arg Ala Val Thr Val Thr Asn Asp Leu Leu 865 870 875 880	2640
acc gac gag gcc gcc tac gct ggg cgc gtg cag gcc gag tat gac acc Thr Asp Glu Ala Ala Tyr Ala Gly Arg Val Gln Gly Glu Tyr Asp Thr 885 890 895	2688
ttg cgc gag cgc cac gcc gag cgg cag gtg cgg ctg att gcg ctg gca Leu Arg Glu Arg His Gly Glu Arg Gln Val Arg Leu Ile Ala Leu Ala 900 905 910	2736
gaa gcc cgc gcc cgc gcc ccg caa ctg agt gcc gcc gtg ccc ccc gcg Glu Ala Arg Ala Arg Ala Pro Gln Leu Ser Ala Ala Val Pro Pro Ala 915 920 925	2784
ccg cac gat ctg ggc cgt cag gtg gtc gaa cag ccc att gcc gag ctg Pro His Asp Leu Gly Arg Gln Val Val Glu Gln Pro Ile Ala Glu Leu 930 935 940	2832
ctg ccc ttc atc gac tgg acg ccc ttt ttc atc gcc tgg gag atg aag Leu Pro Phe Ile Asp Trp Thr Pro Phe Phe Ile Ala Trp Glu Met Lys 945 950 955 960	2880
ggc atc tac ccg ggc atc ctg acc gac cct ctg cgt gcc gag gag gcc Gly Ile Tyr Pro Gly Ile Leu Thr Asp Pro Leu Arg Gly Glu Glu Ala 965 970 975	2928
cgc aag ctg ttt gcc gac gcg cag gcg ctg ctg gag cag gtt atc gcc Arg Lys Leu Phe Ala Asp Ala Gln Ala Leu Leu Glu Gln Val Ile Ala 980 985 990	2976
gac gcc tcg ctg cgg gcg cgc gcc gtc atc ggg ctg tgg ccc gcg cac Asp Gly Ser Leu Arg Ala Arg Gly Val Ile Gly Leu Trp Pro Ala His 995 1000 1005	3024
ggc gac gac atc gtg ctg gac gat gcg gcg atg ggg cgt gcc gag acg Gly Asp Asp Ile Val Leu Asp Asp Ala Ala Met Gly Arg Gly Glu Thr 1010 1015 1020	3072
ctg gat ttc gag acg cac gaa ctc gcc gcc ggg cgc gag ccg ctg ccg Leu Asp Phe Glu Thr His Glu Leu Ala Ala Gly Arg Glu Pro Leu Pro 1025 1030 1035 1040	3120
aac atg ccg cgc ctg cac acg ctg cgg cag cag cgc gac cag acc acg Asn Met Pro Arg Leu His Thr Leu Arg Gln Gln Arg Asp Gln Thr Thr 1045 1050 1055	3168
ccg aac act gcg ctg gct gac ttt gtg gcg gaa gga gcc gac cac atc Pro Asn Thr Ala Leu Ala Asp Phe Val Ala Glu Gly Gly Asp His Ile 1060 1065 1070	3216
ggc gcc ttc gcc acg gcc atc ttc gcc gcc gag gag ttg gcg cag cag Gly Ala Phe Ala Thr Ala Ile Phe Gly Ala Glu Glu Leu Ala Gln Gln 1075 1080 1085	3264
ttc gag gcg cag cac gac gac tac aac tcg att ctg gtc aag gcg gtg Phe Glu Ala Gln His Asp Asp Tyr Asn Ser Ile Leu Val Lys Ala Val 1090 1095 1100	3312
gcc gac cga ctg gcc gag gcc ttt gcc gag aag ctg cac cgc gac gtg	3360

157

Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Lys Leu His Arg Asp Val  
 1105 1110 1115 1120

cgc gtg cgg cac tgg ggt tac gcc gag ggc gag gcg ctc gac aac acc 3408  
 Arg Val Arg His Trp Gly Tyr Ala Glu Gly Glu Ala Leu Asp Asn Thr  
 1125 1130 1135

gac ctc atc aag gag cgc tat cag ggc atc cgc cct gcg ccc ggc tac 3456  
 Asp Leu Ile Lys Glu Arg Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr  
 1140 1145 1150

ccc gcg cag ccc gac cac acc gag aaa cgc acc ctg ttt gag ctg ctg 3504  
 Pro Ala Gln Pro Asp His Thr Glu Lys Arg Thr Leu Phe Glu Leu Leu  
 1155 1160 1165

gac gcg gaa agc atc ggc ctg cgc ctc acc gag tcg tgt gcc atg acc 3552  
 Asp Ala Glu Ser Ile Gly Leu Arg Leu Thr Glu Ser Cys Ala Met Thr  
 1170 1175 1180

ccg gcg gcg gcg gtg tcg ggg ctg tac ttc gcg cat ccg gag gcc cgt 3600  
 Pro Ala Ala Ala Val Ser Gly Leu Tyr Phe Ala His Pro Glu Ala Arg  
 1185 1190 1195 1200

tat ttc gca gtg ggc cgc atc ggg cgc gac cag gtg gag aac tac gcc 3648  
 Tyr Phe Ala Val Gly Arg Ile Gly Arg Asp Gln Val Glu Asn Tyr Ala  
 1205 1210 1215

gcc cgt aag ggt tgg act gtg cag gaa gcc gag cgc tgg ctg ggg ccg 3696  
 Ala Arg Lys Gly Trp Thr Val Gln Glu Ala Glu Arg Trp Leu Gly Pro  
 1220 1225 1230

ctg ctg gcg tac agc gcc ggg ccg ggg cca gaa gca agc cag aaa gcc 3744  
 Leu Leu Ala Tyr Ser Ala Gly Pro Gly Pro Glu Ala Ser Gln Lys Ala  
 1235 1240 1245

ctc ggc gca gag ctg aca gga gcg caa tcg tga 3777  
 Leu Gly Ala Glu Leu Thr Gly Ala Gln Ser  
 1250 1255

&lt;210&gt; 36

&lt;211&gt; 1258

&lt;212&gt; PRT

&lt;213&gt; Deinococcus radiodurans

&lt;400&gt; 36

Met Ser His His Pro Glu Ala Ser Ala Ser Ala Asn Pro Ser Ile Asn  
 1 5 10 15

His Gln Pro Ser Thr Ile Thr Glu Ala Ala Arg Gln Arg Ile Leu Ile  
 20 25 30

Leu Asp Gly Ala Trp Gly Thr Gln Leu Gln Arg Ala Asn Leu Thr Glu  
 35 40 45

Ala Asp Phe Arg Trp Asp Glu Ala Asp Pro Thr Arg Met Tyr Arg Gly  
 50 55 60

Asn Phe Asp Leu Leu Gln Leu Thr Lys Pro Asp Val Ile Arg Ala Val  
 65 70 75 80

His Arg Ala Tyr Phe Glu Ala Gly Ala Asp Ile Ala Ser Thr Asn Thr

158

85	90	95
Phe Asn Ser Thr Thr Ile Ser Gln Ala Asp Tyr Gly Thr Glu Ala Leu 100 105 110		
Ala Tyr Ala Met Asn Arg Glu Gly Ala Arg Leu Ala Arg Glu Val Ala 115 120 125		
Asp Glu Phe Glu Ala Arg Asp Gly Lys Lys Arg Trp Val Ala Gly Ser 130 135 140		
Val Gly Pro Thr Asn Arg Thr Ala Thr Leu Ser Pro Asp Val Glu Arg 145 150 155 160		
Pro Glu Phe Arg Asn Val Thr Tyr Asp Asp Leu Val Ala Ala Tyr Ser 165 170 175		
Glu Ala Ile Thr Gly Leu Met Glu Gly Gly Ala Asp Leu Leu Leu Ile 180 185 190		
Glu Thr Val Phe Asp Thr Leu Asn Ala Lys Ala Ala Leu Phe Ala Ala 195 200 205		
Gln Asp Val Phe Ala Ala Gln Gly Arg Glu Leu Pro Val Met Leu Ser 210 215 220		
Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Pro 225 230 235 240		
Glu Ala Phe Ala Val Ser Thr Glu His Ala Gly Leu Phe Ser Leu Gly 245 250 255		
Leu Asn Cys Ala Leu Gly Ala Asp Leu Leu Arg Pro His Leu Arg Ala 260 265 270		
Ile Ala Ala Asn Thr Glu Ala Leu Val Ser Val His Pro Asn Ala Gly 275 280 285		
Leu Pro Asn Ala Phe Gly Glu Tyr Asp Glu Thr Pro Glu His Thr Ala 290 295 300		
Ala Val Leu Ala Asp Phe Ala Arg Glu Gly Leu Val Asn Ile Val Gly 305 310 315 320		
Gly Cys Cys Gly Thr Thr Pro Glu His Ile Lys Ala Ile Ala Glu Ala 325 330 335		
Val Lys Asp Ile Pro Pro Arg Gln Ala Leu Gln Leu Pro Pro Tyr Leu 340 345 350		
Arg Leu Ser Gly Leu Glu Ala Phe Thr Leu Thr Pro Glu Thr Asn Phe 355 360 365		
Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Pro Lys Phe Ser 370 375 380		
Lys Ala Ile Leu Ala Gly Asp Tyr Asp Ala Gly Leu Lys Ile Ala Arg 385 390 395 400		
Gln Gln Val Thr Asn Gly Ala Gln Ile Val Asp Ile Asn Phe Asp Glu 405 410 415		

159

Gly Met Leu Asp Gly Glu Gly Ala Met Val Lys Phe Leu Asn Leu Leu  
 420 425 430  
 Ala Gly Glu Pro Asp Ile Ser Arg Val Pro Leu Met Leu Asp Ser Ser  
 435 440 445  
 Lys Trp Glu Ile Leu Glu Ala Gly Leu Arg Arg Val Gln Gly Lys Ala  
 450 455 460  
 Val Val Asn Ser Ile Ser Leu Lys Asp Gly Glu Ala Arg Phe Leu Glu  
 465 470 475 480  
 Arg Ala Arg Leu Leu Arg Arg Tyr Gly Ala Ala Val Val Met Ala  
 485 490 495  
 Phe Asp Glu Gln Gly Gln Ala Asp Asn Leu Ala Arg Arg Arg Glu Ile  
 500 505 510  
 Leu Gly Arg Ala Tyr Arg Leu Leu Thr Glu Gln Ala Asp Phe Pro Pro  
 515 520 525  
 Gln Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val Ala Thr Gly Ile  
 530 535 540  
 Glu Glu His Asp Arg Tyr Ala Leu Asp Phe Ile Glu Ala Thr Arg Trp  
 545 550 555 560  
 Ile Lys Glu Asn Leu Pro Ala Ala Lys Val Ser Gly Gly Ile Ser Asn  
 565 570 575  
 Val Ser Phe Ser Phe Arg Gly Asn Asn His Val Arg Glu Ala Met His  
 580 585 590  
 Ala Val Phe Leu Tyr His Ala Ile Arg Ala Gly Leu Asp Met Gly Ile  
 595 600 605  
 Val Asn Ala Gly Met Leu Ala Val Tyr Glu Asp Ile Glu Pro Glu Leu  
 610 615 620  
 Arg Glu Ala Val Glu Asp Val Ile Leu Ala Arg Arg Pro Asp Ala Thr  
 625 630 635 640  
 Glu Arg Leu Leu Thr Leu Ala Asp Arg Tyr Lys Asp Ile Lys Arg Glu  
 645 650 655  
 Ser Ala Ala Gln Ser Ala Trp Arg Asp Leu Pro Val Gln Glu Arg Leu  
 660 665 670  
 Arg His Ala Leu Val Gln Gly Val Ala Asp His Val Asp Glu Asp Ala  
 675 680 685  
 Glu Ala Ala Tyr Gln Glu Leu Gly Ser Pro Leu Ala Val Ile Glu Gly  
 690 695 700  
 Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala Gly  
 705 710 715 720  
 Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys Lys  
 725 730 735  
 Ala Val Ala Tyr Leu Thr Pro Tyr Leu Glu Ala Glu Lys Ala Glu Ser  
 740 745 750

160

Ser Ser Lys Gly Lys Val Leu Leu Ala Thr Val Lys Gly Asp Val His  
 755 760 765  
 Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Ala Cys Asn Gly Tyr  
 770 775 780  
 Gln Val Thr Asp Leu Gly Val Met Val Pro Gly Glu Lys Ile Leu Asp  
 785 790 795 800  
 Glu Ala Glu Arg Leu Gly Ala Asp Val Ile Gly Leu Ser Gly Leu Ile  
 805 810 815  
 Thr Pro Ser Leu Asp Glu Met Val Asn Val Ala Arg Glu Met Thr Arg  
 820 825 830  
 Arg Gly Val Lys Thr Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg  
 835 840 845  
 Ala His Thr Ala Val Lys Ile Asp Pro Ala Tyr Asp Gly Thr Val Val  
 850 855 860  
 His Val Leu Asp Ala Ser Arg Ala Val Thr Val Thr Asn Asp Leu Leu  
 865 870 875 880  
 Thr Asp Glu Ala Ala Tyr Ala Gly Arg Val Gln Gly Glu Tyr Asp Thr  
 885 890 895  
 Leu Arg Glu Arg His Gly Glu Arg Gln Val Arg Leu Ile Ala Leu Ala  
 900 905 910  
 Glu Ala Arg Ala Arg Ala Pro Gln Leu Ser Ala Ala Val Pro Pro Ala  
 915 920 925  
 Pro His Asp Leu Gly Arg Gln Val Val Glu Gln Pro Ile Ala Glu Leu  
 930 935 940  
 Leu Pro Phe Ile Asp Trp Thr Pro Phe Phe Ile Ala Trp Glu Met Lys  
 945 950 955 960  
 Gly Ile Tyr Pro Gly Ile Leu Thr Asp Pro Leu Arg Gly Glu Glu Ala  
 965 970 975  
 Arg Lys Leu Phe Ala Asp Ala Gln Ala Leu Leu Glu Gln Val Ile Ala  
 980 985 990  
 Asp Gly Ser Leu Arg Ala Arg Gly Val Ile Gly Leu Trp Pro Ala His  
 995 1000 1005  
 Gly Asp Asp Ile Val Leu Asp Asp Ala Ala Met Gly Arg Gly Glu Thr  
 1010 1015 1020  
 Leu Asp Phe Glu Thr His Glu Leu Ala Ala Gly Arg Glu Pro Leu Pro  
 1025 1030 1035 1040  
 Asn Met Pro Arg Leu His Thr Leu Arg Gln Gln Arg Asp Gln Thr Thr  
 1045 1050 1055  
 Pro Asn Thr Ala Leu Ala Asp Phe Val Ala Glu Gly Gly Asp His Ile  
 1060 1065 1070  
 Gly Ala Phe Ala Thr Ala Ile Phe Gly Ala Glu Glu Leu Ala Gln Gln

## 161

1075                      1080                      1085

Phe Glu Ala Gln His Asp Asp Tyr Asn Ser Ile Leu Val Lys Ala Val  
1090                      1095                      1100

Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Lys Leu His Arg Asp Val  
1105                      1110                      1115                      1120

Arg Val Arg His Trp Gly Tyr Ala Glu Gly Glu Ala Leu Asp Asn Thr  
1125                      1130                      1135

Asp Leu Ile Lys Glu Arg Tyr Gln Gly Ile Arg Pro Ala Pro Gly Tyr  
1140                      1145                      1150

Pro Ala Gln Pro Asp His Thr Glu Lys Arg Thr Leu Phe Glu Leu Leu  
1155                      1160                      1165

Asp Ala Glu Ser Ile Gly Leu Arg Leu Thr Glu Ser Cys Ala Met Thr  
1170                      1175                      1180

Pro Ala Ala Ala Val Ser Gly Leu Tyr Phe Ala His Pro Glu Ala Arg  
1185                      1190                      1195                      1200

Tyr Phe Ala Val Gly Arg Ile Gly Arg Asp Gln Val Glu Asn Tyr Ala  
1205                      1210                      1215

Ala Arg Lys Gly Trp Thr Val Gln Glu Ala Glu Arg Trp Leu Gly Pro  
1220                      1225                      1230

Leu Leu Ala Tyr Ser Ala Gly Pro Gly Pro Glu Ala Ser Gln Lys Ala  
1235                      1240                      1245

Leu Gly Ala Glu Leu Thr Gly Ala Gln Ser  
1250                      1255

<210> 37  
<211> 3642  
<212> DNA  
<213> Clostridium acetobutylicum

<220>  
<221> CDS  
<222> (1) .. (3639)  
<223> RCA01265

<400> 37

ctt atg aat tct tca cta aag aat ttg tta aat aac aaa att tta gtt 48  
Leu Met Asn Ser Ser Leu Lys Asn Leu Leu Asn Asn Lys Ile Leu Val  
1                      5                      10                      15

tta gat ggt gct atg gga aca tgt att caa tcc ttt aat cta gat gaa 96  
Leu Asp Gly Ala Met Gly Thr Cys Ile Gln Ser Phe Asn Leu Asp Glu  
20                      25                      30

ggc gac ttt aaa ggt tcc tta tct tgt aca tgt cat tcc aat caa aaa 144  
Gly Asp Phe Lys Gly Ser Leu Ser Cys Thr Cys His Ser Asn Gln Lys  
35                      40                      45

gga aac aat gat gtt tta aat tta acc aag cca gaa ata ata aaa gaa 192  
Gly Asn Asn Asp Val Leu Asn Leu Thr Lys Pro Glu Ile Ile Lys Glu  
50                      55                      60

162

atc cac aag aga tac ctt gaa gct ggc gca gat ata ata gaa aca aac Ile His Lys Arg Tyr Leu Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn 65 70 75 80	240
act ttt aac gct act gaa ata tca caa aaa gat tat aat atg caa gat Thr Phe Asn Ala Thr Glu Ile Ser Gln Lys Asp Tyr Asn Met Gln Asp 85 90 95	288
aaa ata tat gat att aat ttt aag ggg gca aaa ctc gca aag gaa gct Lys Ile Tyr Asp Ile Asn Phe Lys Gly Ala Lys Leu Ala Lys Glu Ala 100 105 110	336
tgt act tac tac aca aaa cta aat cct aat aag cct aga ttt gct gct Cys Thr Tyr Tyr Thr Lys Leu Asn Pro Asn Lys Pro Arg Phe Ala Ala 115 120 125	384
ggg tct att ggg cct aca aat aga act gct tct cta tct cca gat gtt Gly Ser Ile Gly Pro Thr Asn Arg Thr Ala Ser Leu Ser Pro Asp Val 130 135 140	432
gaa aat cct ggt ttt aga aat gta acc ttt gat gag cta tgt aat gcc Glu Asn Pro Gly Phe Arg Asn Val Thr Phe Asp Glu Leu Cys Asn Ala 145 150 155 160	480
tat aaa cat caa ata gag gct cta ata gat gga ggt gta gac ctt ctt Tyr Lys His Gln Ile Glu Ala Leu Ile Asp Gly Gly Val Asp Leu Leu 165 170 175	528
tta att gaa act ata ttt gat act tta aac gct aga gca gca atc ttt Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Arg Ala Ala Ile Phe 180 185 190	576
gca gca gaa aca gta ttt gaa aat aaa aaa ata aaa ctt cct att ata Ala Ala Glu Thr Val Phe Glu Asn Lys Lys Ile Lys Leu Pro Ile Ile 195 200 205	624
att tca ggg aca ata gct gat aaa agt gga aga ata tta tcc ggt caa Ile Ser Gly Thr Ile Ala Asp Lys Ser Gly Arg Ile Leu Ser Gly Gln 210 215 220	672
act ctt gac gct ttt gca gaa agt tta aaa aac gaa aat ata att gct Thr Leu Asp Ala Phe Ala Glu Ser Leu Lys Asn Glu Asn Ile Ile Ala 225 230 235 240	720
ata ggg ctt aat tgt tcc ttt ggt gct gaa gaa ctt ata cct ttt ata Ile Gly Leu Asn Cys Ser Phe Gly Ala Glu Glu Leu Ile Pro Phe Ile 245 250 255	768
aaa aga ctc tct gaa aca caa aat aga tat ata tcc ttt cat cca aac Lys Arg Leu Ser Glu Thr Gln Asn Arg Tyr Ile Ser Phe His Pro Asn 260 265 270	816
gca gga ctt cca aac tcc ctt ggt gaa tat gaa gaa ctg cca gag gaa Ala Gly Leu Pro Asn Ser Leu Gly Glu Tyr Glu Glu Leu Pro Glu Glu 275 280 285	864
act gct agc att gta aaa aaa tta gca ctt gaa gga cat tta aat ata Thr Ala Ser Ile Val Lys Lys Leu Ala Leu Glu Gly His Leu Asn Ile 290 295 300	912
gtt gga ggc tgc tgt ggc act aca cca gaa cat ata aga gca ata agc	960



163																		
Val	Gly	Gly	Cys	Cys	Gly	Thr	Thr	Pro	Glu	His	Ile	Arg	Ala	Ile	Ser			
305					310					315					320			
agc	gta	gtt	aaa	ggc	att	tct	cca	aga	aaa	gtt	cca	aac	ttg	gaa	ccc	1008		
Ser	Val	Val	Lys	Gly	Ile	Ser	Pro	Arg	Lys	Val	Pro	Asn	Leu	Glu	Pro			
			325						330					335				
aaa	aca	att	tac	agc	gga	cta	gaa	aac	ata	aaa	att	gat	aag	aac	agt	1056		
Lys	Thr	Ile	Tyr	Ser	Gly	Leu	Glu	Asn	Ile	Lys	Ile	Asp	Lys	Asn	Ser			
			340					345					350					
aac	ttc	ata	aat	ata	ggc	gaa	aga	aca	aat	gta	gcg	ggc	tca	aga	aaa	1104		
Asn	Phe	Ile	Asn	Ile	Gly	Glu	Arg	Thr	Asn	Val	Ala	Gly	Ser	Arg	Lys			
			355				360					365						
ttc	gca	agg	ctt	ata	cgt	gaa	aaa	aat	tat	gag	gag	gct	cta	acc	att	1152		
Phe	Ala	Arg	Leu	Ile	Arg	Glu	Lys	Asn	Tyr	Glu	Glu	Ala	Leu	Thr	Ile			
			370			375					380							
gca	aga	cat	cag	gtt	gaa	aat	ggt	gcc	caa	att	ata	gat	ata	aat	ttt	1200		
Ala	Arg	His	Gln	Val	Glu	Asn	Gly	Ala	Gln	Ile	Ile	Asp	Ile	Asn	Phe			
385					390					395					400			
gat	gat	gca	ctt	tta	gat	gct	cgc	tct	gaa	atg	gaa	aca	ttt	tta	aga	1248		
Asp	Asp	Ala	Leu	Leu	Asp	Ala	Arg	Ser	Glu	Met	Glu	Thr	Phe	Leu	Arg			
				405					410					415				
ctt	att	gca	agt	gaa	cct	gaa	ata	tca	aaa	gtt	cca	gtt	atg	ata	gac	1296		
Leu	Ile	Ala	Ser	Glu	Pro	Glu	Ile	Ser	Lys	Val	Pro	Val	Met	Ile	Asp			
			420				425						430					
tcc	tct	aat	ttt	gaa	gtt	tta	aaa	gtt	gga	tta	aag	tct	att	caa	ggt	1344		
Ser	Ser	Asn	Phe	Glu	Val	Leu	Lys	Val	Gly	Leu	Lys	Ser	Ile	Gln	Gly			
		435					440					445						
aaa	gcc	ata	gta	aat	tcc	ata	agt	ctt	aag	gtt	gga	gaa	gaa	aag	ttc	1392		
Lys	Ala	Ile	Val	Asn	Ser	Ile	Ser	Leu	Lys	Val	Gly	Glu	Glu	Lys	Phe			
		450				455					460							
att	gaa	gag	gca	aaa	ttt	ata	aag	aac	ttt	ggc	gct	ggc	gta	gtt	gta	1440		
Ile	Glu	Glu	Ala	Lys	Phe	Ile	Lys	Asn	Phe	Gly	Ala	Gly	Val	Val	Val			
465					470					475				480				
atg	gcc	ttt	gac	gaa	gaa	ggt	caa	gca	gct	act	tat	gaa	aga	aaa	att	1488		
Met	Ala	Phe	Asp	Glu	Glu	Gly	Gln	Ala	Ala	Thr	Tyr	Glu	Arg	Lys	Ile			
				485					490					495				
gaa	atc	tgc	aag	aga	gct	tat	act	att	ctc	aca	gaa	aaa	gtt	gag	ttt	1536		
Glu	Ile	Cys	Lys	Arg	Ala	Tyr	Thr	Ile	Le									

164

agc aac ctc tcc ttt tct ttt agg ggt aat gac gca ata aga aga gct 1728  
 Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Asp Ala Ile Arg Arg Ala  
 565 570 575

atg cat tct gtt ttc ctt tac cat gca ata aac gct gga atg gat atg 1776  
 Met His Ser Val Phe Leu Tyr His Ala Ile Asn Ala Gly Met Asp Met  
 580 585 590

ggt att gtt aat cca gca atg att gat tta tat gac gat ata gat aag 1824  
 Gly Ile Val Asn Pro Ala Met Ile Asp Leu Tyr Asp Asp Ile Asp Lys  
 595 600 605

gat ctt ctc gaa aag gtt gag aat gtt gta cta aat aaa tca tct aac 1872  
 Asp Leu Leu Glu Lys Val Glu Asn Val Val Leu Asn Lys Ser Ser Asn  
 610 615 620

gct tct gaa tca tta cta gaa ttt gct caa acg tat aaa aag acg act 1920  
 Ala Ser Glu Ser Leu Leu Glu Phe Ala Gln Thr Tyr Lys Lys Thr Thr  
 625 630 635 640

gaa acc tta gaa aag cac gag gat gaa tgg cga caa aaa agc cca agt 1968  
 Glu Thr Leu Glu Lys His Glu Asp Glu Trp Arg Gln Lys Ser Pro Ser  
 645 650 655

gaa agg ttg agt tat gct tta gtt aaa gga aat gtt gaa ttt att gaa 2016  
 Glu Arg Leu Ser Tyr Ala Leu Val Lys Gly Asn Val Glu Phe Ile Glu  
 660 665 670

gaa gat ata gaa gaa gca aga aaa gag tat aca aat gca ctt gaa att 2064  
 Glu Asp Ile Glu Glu Ala Arg Lys Glu Tyr Thr Asn Ala Leu Glu Ile  
 675 680 685

ata gag gtt cct tta atg aat gga atg aaa aaa gtg ggt aaa ctt ttt 2112  
 Ile Glu Val Pro Leu Met Asn Gly Met Lys Lys Val Gly Lys Leu Phe  
 690 695 700

gga gag gga aaa atg ttt ctt cct caa gta gta aaa agt gct aga gtt 2160  
 Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val  
 705 710 715 720

atg aaa aag gct gtt gaa tgt ctt ctt ccc tat ata aac gaa gaa aag 2208  
 Met Lys Lys Ala Val Glu Cys Leu Leu Pro Tyr Ile Asn Glu Glu Lys  
 725 730 735

tct aaa aat cac aat aaa agt gct ggt aag gtt gta ttt gca act gtt 2256  
 Ser Lys Asn His Asn Lys Ser Ala Gly Lys Val Val Phe Ala Thr Val  
 740 745 750

aaa ggc gat gtt cat gac ata ggc aaa aat atc gta tct gta gtt ctt 2304  
 Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser Val Val Leu  
 755 760 765

tcc tgc aac aat ttt gaa gtt ata gat tta gga gta atg gtt ccc cct 2352  
 Ser Cys Asn Asn Phe Glu Val Ile Asp Leu Gly Val Met Val Pro Pro  
 770 775 780

gaa acc ata ctt gaa acg gca aaa cgt gaa aat gca gat atc att gct 2400  
 Glu Thr Ile Leu Glu Thr Ala Lys Arg Glu Asn Ala Asp Ile Ile Ala  
 785 790 795 800

tta agt ggt tta att aca cct tct ctt aat gaa atg gct tat gta gct 2448

165

Leu Ser Gly Leu Ile Thr Pro Ser Leu Asn Glu Met Ala Tyr Val Ala	
805 810 815	
gaa gaa atg aaa agg ctt aat ttt gat ata cca ctt atg gtg ggt ggt	2496
Glu Glu Met Lys Arg Leu Asn Phe Asp Ile Pro Leu Met Val Gly Gly	
820 825 830	
gct gct acc tca aaa act cac aca gct tta aaa cta gct acg aaa tat	2544
Ala Ala Thr Ser Lys Thr His Thr Ala Leu Lys Leu Ala Thr Lys Tyr	
835 840 845	
aaa tat gta gta cac agt act gat gct tca gat gct gtt acc gta gcc	2592
Lys Tyr Val Val His Ser Thr Asp Ala Ser Asp Ala Val Thr Val Ala	
850 855 860	
aaa aat cta atg agt gaa aac aaa ttt act ttc tta gaa aaa tta aat	2640
Lys Asn Leu Met Ser Glu Asn Lys Phe Thr Phe Leu Glu Lys Leu Asn	
865 870 875 880	
gaa gag tat tct aaa ata aga gag acc ttc tct act aat aag att gaa	2688
Glu Glu Tyr Ser Lys Ile Arg Glu Thr Phe Ser Thr Asn Lys Ile Glu	
885 890 895	
ctt atc tcc att caa aac gca aga aaa aac aga ttt act att gac tgg	2736
Leu Ile Ser Ile Gln Asn Ala Arg Lys Asn Arg Phe Thr Ile Asp Trp	
900 905 910	
aat aaa act aaa ata act gaa cct aaa ttt gtc ggt ata aaa aaa tta	2784
Asn Lys Thr Lys Ile Thr Glu Pro Lys Phe Val Gly Ile Lys Lys Leu	
915 920 925	
caa gct gta cct ata aat gaa tta aga aag tat ata gat tgg act ttc	2832
Gln Ala Val Pro Ile Asn Glu Leu Arg Lys Tyr Ile Asp Trp Thr Phe	
930 935 940	
ttc ttt acg tct tgg gat atg gga atg aat tac ccc aaa ata atg aaa	2880
Phe Phe Thr Ser Trp Asp Met Gly Met Asn Tyr Pro Lys Ile Met Lys	
945 950 955 960	
gat cct aaa tac gga gct gaa gct caa aaa ctc ttt aag gat gcc aat	2928
Asp Pro Lys Tyr Gly Ala Glu Ala Gln Lys Leu Phe Lys Asp Ala Asn	
965 970 975	
gaa atg ctt gat tta ttg caa aaa gaa aat tta atc act tgt aat gga	2976
Glu Met Leu Asp Leu Leu Gln Lys Glu Asn Leu Ile Thr Cys Asn Gly	
980 985 990	
gtt ttt gga ata ttc cca gct aat tct gtt aat gat gat ata gaa atc	3024
Val Phe Gly Ile Phe Pro Ala Asn Ser Val Asn Asp Asp Ile Glu Ile	
995 1000 1005	
tac act gat aaa gga act gta acc ata aat act ctt cgt cag cag cag	3072
Tyr Thr Asp Lys Gly Thr Val Thr Ile Asn Thr Leu Arg Gln Gln Gln	
1010 1015 1020	
ata ctt aaa gac agc gat tat aaa gct cta tct gat tat atc gct cca	3120
Ile Leu Lys Asp Ser Asp Tyr Lys Ala Leu Ser Asp Tyr Ile Ala Pro	
1025 1030 1035 1040	
aag ggt att ggc atc aaa gat tat ata ggt ggt ttt att gta act gct	3168
Lys Gly Ile Gly Ile Lys Asp Tyr Ile Gly Gly Phe Ile Val Thr Ala	
1045 1050 1055	

166

gga ata ggt gca aag gaa tat tcc gat aaa tta aag aaa aaa tgc gac 3216  
 Gly Ile Gly Ala Lys Glu Tyr Ser Asp Lys Leu Lys Lys Lys Cys Asp  
 1060 1065 1070

gat tat gga gct act atg ctt aaa ctt ata tgc gat aga ctt gca gag 3264  
 Asp Tyr Gly Ala Thr Met Leu Lys Leu Ile Cys Asp Arg Leu Ala Glu  
 1075 1080 1085

gcc ttt tca gaa ctt ctt cac cta agg gta aga aaa gaa tac tgg gga 3312  
 Ala Phe Ser Glu Leu Leu His Leu Arg Val Arg Lys Glu Tyr Trp Gly  
 1090 1095 1100

tac tct caa gat gaa aac tta tcc tta gaa aaa ctt ctt aaa gga agt 3360  
 Tyr Ser Gln Asp Glu Asn Leu Ser Leu Glu Lys Leu Leu Lys Gly Ser  
 1105 1110 1115 1120

tac aga ggg ata aaa cca gct att gga tat cct tct att ccc gat cac 3408  
 Tyr Arg Gly Ile Lys Pro Ala Ile Gly Tyr Pro Ser Ile Pro Asp His  
 1125 1130 1135

tct gaa aaa gca aag tta ttt gat tta ctt tta ggt aaa act tca ata 3456  
 Ser Glu Lys Ala Lys Leu Phe Asp Leu Leu Leu Gly Lys Thr Ser Ile  
 1140 1145 1150

gga gtg gaa ttg acg gaa agt tat atg atg aat cca act tca agt gta 3504  
 Gly Val Glu Leu Thr Glu Ser Tyr Met Met Asn Pro Thr Ser Ser Val  
 1155 1160 1165

tgc ggt ttg tat ttt gca aat gaa cga gca aaa tac ttt aat ata aat 3552  
 Cys Gly Leu Tyr Phe Ala Asn Glu Arg Ala Lys Tyr Phe Asn Ile Asn  
 1170 1175 1180

aaa ata gga aaa gat caa ctt gag gac tat gct gtt cga agt aat aaa 3600  
 Lys Ile Gly Lys Asp Gln Leu Glu Asp Tyr Ala Val Arg Ser Asn Lys  
 1185 1190 1195 1200

gac att aat gaa ata aaa aaa tta tta gat act ctg tta taa 3642  
 Asp Ile Asn Glu Ile Lys Lys Leu Leu Asp Thr Leu Leu  
 1205 1210

<210> 38  
 <211> 1213  
 <212> PRT  
 <213> Clostridium acetobutylicum

<400> 38  
 Leu Met Asn Ser Ser Leu Lys Asn Leu Leu Asn Asn Lys Ile Leu Val  
 1 5 10 15  
 Leu Asp Gly Ala Met Gly Thr Cys Ile Gln Ser Phe Asn Leu Asp Glu  
 20 25 30  
 Gly Asp Phe Lys Gly Ser Leu Ser Cys Thr Cys His Ser Asn Gln Lys  
 35 40 45  
 Gly Asn Asn Asp Val Leu Asn Leu Thr Lys Pro Glu Ile Ile Lys Glu  
 50 55 60  
 Ile His Lys Arg Tyr Leu Glu Ala Gly Ala Asp Ile Ile Glu Thr Asn  
 65 70 75 80

167

Thr Phe Asn Ala Thr Glu Ile Ser Gln Lys Asp Tyr Asn Met Gln Asp  
 85 90 95  
 Lys Ile Tyr Asp Ile Asn Phe Lys Gly Ala Lys Leu Ala Lys Glu Ala  
 100 105 110  
 Cys Thr Tyr Tyr Thr Lys Leu Asn Pro Asn Lys Pro Arg Phe Ala Ala  
 115 120 125  
 Gly Ser Ile Gly Pro Thr Asn Arg Thr Ala Ser Leu Ser Pro Asp Val  
 130 135 140  
 Glu Asn Pro Gly Phe Arg Asn Val Thr Phe Asp Glu Leu Cys Asn Ala  
 145 150 155 160  
 Tyr Lys His Gln Ile Glu Ala Leu Ile Asp Gly Gly Val Asp Leu Leu  
 165 170 175  
 Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Arg Ala Ala Ile Phe  
 180 185 190  
 Ala Ala Glu Thr Val Phe Glu Asn Lys Lys Ile Lys Leu Pro Ile Ile  
 195 200 205  
 Ile Ser Gly Thr Ile Ala Asp Lys Ser Gly Arg Ile Leu Ser Gly Gln  
 210 215 220  
 Thr Leu Asp Ala Phe Ala Glu Ser Leu Lys Asn Glu Asn Ile Ile Ala  
 225 230 235 240  
 Ile Gly Leu Asn Cys Ser Phe Gly Ala Glu Glu Leu Ile Pro Phe Ile  
 245 250 255  
 Lys Arg Leu Ser Glu Thr Gln Asn Arg Tyr Ile Ser Phe His Pro Asn  
 260 265 270  
 Ala Gly Leu Pro Asn Ser Leu Gly Glu Tyr Glu Glu Leu Pro Glu Glu  
 275 280 285  
 Thr Ala Ser Ile Val Lys Lys Leu Ala Leu Glu Gly His Leu Asn Ile  
 290 295 300  
 Val Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile Arg Ala Ile Ser  
 305 310 315 320  
 Ser Val Val Lys Gly Ile Ser Pro Arg Lys Val Pro Asn Leu Glu Pro  
 325 330 335  
 Lys Thr Ile Tyr Ser Gly Leu Glu Asn Ile Lys Ile Asp Lys Asn Ser  
 340 345 350  
 Asn Phe Ile Asn Ile Gly Glu Arg Thr Asn Val Ala Gly Ser Arg Lys  
 355 360 365  
 Phe Ala Arg Leu Ile Arg Glu Lys Asn Tyr Glu Glu Ala Leu Thr Ile  
 370 375 380  
 Ala Arg His Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile Asn Phe  
 385 390 395 400  
 Asp Asp Ala Leu Leu Asp Ala Arg Ser Glu Met Glu Thr Phe Leu Arg

168

405                      410                      415

Leu Ile Ala Ser Glu Pro Glu Ile Ser Lys Val Pro Val Met Ile Asp  
420                      425                      430

Ser Ser Asn Phe Glu Val Leu Lys Val Gly Leu Lys Ser Ile Gln Gly  
435                      440                      445

Lys Ala Ile Val Asn Ser Ile Ser Leu Lys Val Gly Glu Glu Lys Phe  
450                      455                      460

Ile Glu Glu Ala Lys Phe Ile Lys Asn Phe Gly Ala Gly Val Val Val  
465                      470                      475                      480

Met Ala Phe Asp Glu Glu Gly Gln Ala Ala Thr Tyr Glu Arg Lys Ile  
485                      490                      495

Glu Ile Cys Lys Arg Ala Tyr Thr Ile Leu Thr Glu Lys Val Glu Phe  
500                      505                      510

Pro Pro Glu Asn Ile Ile Phe Asp Pro Asn Ile Leu Ser Ile Ala Thr  
515                      520                      525

Gly Ile Glu Glu His Asp Asn Tyr Ala Val Asn Tyr Ile Lys Ala Val  
530                      535                      540

Lys Trp Ile Lys Glu Asn Leu Pro Tyr Ala Lys Val Ser Gly Gly Val  
545                      550                      555                      560

Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Asp Ala Ile Arg Arg Ala  
565                      570                      575

Met His Ser Val Phe Leu Tyr His Ala Ile Asn Ala Gly Met Asp Met  
580                      585                      590

Gly Ile Val Asn Pro Ala Met Ile Asp Leu Tyr Asp Asp Ile Asp Lys  
595                      600                      605

Asp Leu Leu Glu Lys Val Glu Asn Val Val Leu Asn Lys Ser Ser Asn  
610                      615                      620

Ala Ser Glu Ser Leu Leu Glu Phe Ala Gln Thr Tyr Lys Lys Thr Thr  
625                      630                      635                      640

Glu Thr Leu Glu Lys His Glu Asp Glu Trp Arg Gln Lys Ser Pro Ser  
645                      650                      655

Glu Arg Leu Ser Tyr Ala Leu Val Lys Gly Asn Val Glu Phe Ile Glu  
660                      665                      670

Glu Asp Ile Glu Glu Ala Arg Lys Glu Tyr Thr Asn Ala Leu Glu Ile  
675                      680                      685

Ile Glu Val Pro Leu Met Asn Gly Met Lys Lys Val Gly Lys Leu Phe  
690                      695                      700

Gly Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val  
705                      710                      715                      720

Met Lys Lys Ala Val Glu Cys Leu Leu Pro Tyr Ile Asn Glu Glu Lys  
725                      730                      735

169

Ser Lys Asn His Asn Lys Ser Ala Gly Lys Val Val Phe Ala Thr Val  
 740 745 750  
 Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser Val Val Leu  
 755 760 765  
 Ser Cys Asn Asn Phe Glu Val Ile Asp Leu Gly Val Met Val Pro Pro  
 770 775 780  
 Glu Thr Ile Leu Glu Thr Ala Lys Arg Glu Asn Ala Asp Ile Ile Ala  
 785 790 795 800  
 Leu Ser Gly Leu Ile Thr Pro Ser Leu Asn Glu Met Ala Tyr Val Ala  
 805 810 815  
 Glu Glu Met Lys Arg Leu Asn Phe Asp Ile Pro Leu Met Val Gly Gly  
 820 825 830  
 Ala Ala Thr Ser Lys Thr His Thr Ala Leu Lys Leu Ala Thr Lys Tyr  
 835 840 845  
 Lys Tyr Val Val His Ser Thr Asp Ala Ser Asp Ala Val Thr Val Ala  
 850 855 860  
 Lys Asn Leu Met Ser Glu Asn Lys Phe Thr Phe Leu Glu Lys Leu Asn  
 865 870 875 880  
 Glu Glu Tyr Ser Lys Ile Arg Glu Thr Phe Ser Thr Asn Lys Ile Glu  
 885 890 895  
 Leu Ile Ser Ile Gln Asn Ala Arg Lys Asn Arg Phe Thr Ile Asp Trp  
 900 905 910  
 Asn Lys Thr Lys Ile Thr Glu Pro Lys Phe Val Gly Ile Lys Lys Leu  
 915 920 925  
 Gln Ala Val Pro Ile Asn Glu Leu Arg Lys Tyr Ile Asp Trp Thr Phe  
 930 935 940  
 Phe Phe Thr Ser Trp Asp Met Gly Met Asn Tyr Pro Lys Ile Met Lys  
 945 950 955 960  
 Asp Pro Lys Tyr Gly Ala Glu Ala Gln Lys Leu Phe Lys Asp Ala Asn  
 965 970 975  
 Glu Met Leu Asp Leu Leu Gln Lys Glu Asn Leu Ile Thr Cys Asn Gly  
 980 985 990  
 Val Phe Gly Ile Phe Pro Ala Asn Ser Val Asn Asp Asp Ile Glu Ile  
 995 1000 1005  
 Tyr Thr Asp Lys Gly Thr Val Thr Ile Asn Thr Leu Arg Gln Gln Gln  
 1010 1015 1020  
 Ile Leu Lys Asp Ser Asp Tyr Lys Ala Leu Ser Asp Tyr Ile Ala Pro  
 1025 1030 1035 1040  
 Lys Gly Ile Gly Ile Lys Asp Tyr Ile Gly Gly Phe Ile Val Thr Ala  
 1045 1050 1055  
 Gly Ile Gly Ala Lys Glu Tyr Ser Asp Lys Leu Lys Lys Lys Cys Asp  
 1060 1065 1070

170

Asp Tyr Gly Ala Thr Met Leu Lys Leu Ile Cys Asp Arg Leu Ala Glu  
 1075 1080 1085

Ala Phe Ser Glu Leu Leu His Leu Arg Val Arg Lys Glu Tyr Trp Gly  
 1090 1095 1100

Tyr Ser Gln Asp Glu Asn Leu Ser Leu Glu Lys Leu Leu Lys Gly Ser  
 1105 1110 1115 1120

Tyr Arg Gly Ile Lys Pro Ala Ile Gly Tyr Pro Ser Ile Pro Asp His  
 1125 1130 1135

Ser Glu Lys Ala Lys Leu Phe Asp Leu Leu Leu Gly Lys Thr Ser Ile  
 1140 1145 1150

Gly Val Glu Leu Thr Glu Ser Tyr Met Met Asn Pro Thr Ser Ser Val  
 1155 1160 1165

Cys Gly Leu Tyr Phe Ala Asn Glu Arg Ala Lys Tyr Phe Asn Ile Asn  
 1170 1175 1180

Lys Ile Gly Lys Asp Gln Leu Glu Asp Tyr Ala Val Arg Ser Asn Lys  
 1185 1190 1195 1200

Asp Ile Asn Glu Ile Lys Lys Leu Leu Asp Thr Leu Leu  
 1205 1210

&lt;210&gt; 39

&lt;211&gt; 3954

&lt;212&gt; DNA

<213> *Caulobacter\_crescentus*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (3951)

&lt;223&gt; RCO02271

&lt;400&gt; 39

atg acc gat ctc tcc atc cgc gcc aac cgc gtc gcc gcc ctg aag gcc 48  
 Met Thr Asp Leu Ser Ile Arg Ala Asn Arg Val Ala Ala Leu Lys Ala  
 1 5 10 15

gcc gcc aag gag cgt att ctc att ctc gac ggc tcc tgg ggc gtg atg 96  
 Ala Ala Lys Glu Arg Ile Leu Ile Leu Asp Gly Ser Trp Gly Val Met  
 20 25 30

ttc cag aag aag ggg ctg acc gag gcc gac tac cgc gcc gag cgc ttc 144  
 Phe Gln Lys Lys Gly Leu Thr Glu Ala Asp Tyr Arg Ala Glu Arg Phe  
 35 40 45

gcc gcc tac aac ggc cag atg aag ggc aat aac gac atc ctg tgc ctg 192  
 Ala Ala Tyr Asn Gly Gln Met Lys Gly Asn Asn Asp Ile Leu Cys Leu  
 50 55 60

acg cgg ccc gat ctc gtg gcc gag ctg cac gac gcc tat ttc agc gcc 240  
 Thr Arg Pro Asp Leu Val Ala Glu Leu His Asp Ala Tyr Phe Ser Ala  
 65 70 75 80

ggc gcc gac atc tcc gag acc aac acc ttc tgc ggc acc acc atc gcc 288  
 Gly Ala Asp Ile Ser Glu Thr Asn Thr Phe Ser Gly Thr Thr Ile Ala



171

85	90	95	
cag gcc gac tat cat ctg ggt gaa cag gat gtc tgg gac atc aac ctg			336
Gln Ala Asp Tyr His Leu Gly Glu Gln Asp Val Trp Asp Ile Asn Leu			
100	105	110	
gaa ggc gcc aag atc ggc cgc tcg gtg gcc gac cgc tgg aac gcg cag			384
Glu Gly Ala Lys Ile Gly Arg Ser Val Ala Asp Arg Trp Asn Ala Gln			
115	120	125	
aat ccc gac cgc ccg aag ttc atc gcc ggc tcg atg ggg ccg ctg aac			432
Asn Pro Asp Arg Pro Lys Phe Ile Ala Gly Ser Met Gly Pro Leu Asn			
130	135	140	
gtc atg ctg tcg atg tcg tcg gac gtg aac gat ccg ggc gcg cgc aag			480
Val Met Leu Ser Met Ser Ser Asp Val Asn Asp Pro Gly Ala Arg Lys			
145	150	155	160
gtg acc ttc gac cag gtc tac gag gcc tat cgc cag cag gtg gat gcg			528
Val Thr Phe Asp Gln Val Tyr Glu Ala Tyr Arg Gln Gln Val Asp Ala			
165	170	175	
ctt tac cag ggc ggg gtc gat ctc ttc ctg atc gag acc atc acc gac			576
Leu Tyr Gln Gly Gly Val Asp Leu Phe Leu Ile Glu Thr Ile Thr Asp			
180	185	190	
acc ctg aac tgc aag gcc gcg atc aag gcg atc ctg gac tgg cgc gac			624
Thr Leu Asn Cys Lys Ala Ala Ile Lys Ala Ile Leu Asp Trp Arg Asp			
195	200	205	
gag ggc cac gag gag ctg ccg atc tgg atc agc ggc acc atc acc gat			672
Glu Gly His Glu Glu Leu Pro Ile Trp Ile Ser Gly Thr Ile Thr Asp			
210	215	220	
cgc tcg ggc cgc acc ctg tcg ggc cag acg gcc gag gcg ttc tgg aac			720
Arg Ser Gly Arg Thr Leu Ser Gly Gln Thr Ala Glu Ala Phe Trp Asn			
225	230	235	240
agc gtc aag cac gcc aag ccg ttc gca gtg ggc ttc aac tgc gcc ctg			768
Ser Val Lys His Ala Lys Pro Phe Ala Val Gly Phe Asn Cys Ala Leu			
245	250	255	
ggc gcg gat ttg atg cgt ccg cac atc gcc gag atg gcc cgt atc gcc			816
Gly Ala Asp Leu Met Arg Pro His Ile Ala Glu Met Ala Arg Ile Ala			
260	265	270	
gac acc ctg gtc gca gcc tat ccc aac gcc ggc ctg ccc aac gcc atg			864
Asp Thr Leu Val Ala Ala Tyr Pro Asn Ala Gly Leu Pro Asn Ala Met			
275	280	285	
ggc cag tac gac gag gag ccg cac gag acc ggc cac gcc ctg cac gag			912
Gly Gln Tyr Asp Glu Glu Pro His Glu Thr Gly His Ala Leu His Glu			
290	295	300	
tgg gcc aag gac ggc ctc gtc aac atc ctg ggc ggc tgc tgc ggc acg			960
Trp Ala Lys Asp Gly Leu Val Asn Ile Leu Gly Gly Cys Cys Gly Thr			
305	310	315	320
aca ccg gac cac atc cgt cac gtc gcc gac gag gtg cgc ggc gtg acg			1008
Thr Pro Asp His Ile Arg His Val Ala Asp Glu Val Arg Gly Val Thr			
325	330	335	

172

ccg cgc cag atc ccc gag cgc ccc aag gcc atg cgc ctg gcg ggc ctc Pro Arg Gln Ile Pro Glu Arg Pro Lys Ala Met Arg Leu Ala Gly Leu 340 345 350	1056
gaa ccg ttc gag ttg gct tag tgg cta cgg ccg caa att ccc ttc tcc Glu Pro Phe Glu Leu Ala Xaa Trp Leu Arg Pro Gln Ile Pro Phe Ser 355 360 365	1104
cct tgc ggg aga agg tgt cgc cga agg cga cgg atg agg ggt ctc gcc Pro Cys Gly Arg Arg Cys Arg Arg Arg Arg Arg Met Arg Gly Leu Ala 370 375 380	1152
ggc cct tca acc gct gtc tcg cgg cgg cga cgt tct tca acc cct cat Gly Pro Ser Thr Ala Val Ser Arg Arg Arg Ser Ser Thr Pro His 385 390 395 400	1200
ccg acc cgc tgc gcg ggc cac ctt ctc ccg caa ggg gag aag gga tga Pro Thr Arg Cys Ala Gly His Leu Leu Pro Gln Gly Glu Lys Gly Xaa 405 410 415	1248
ctg cta ttg gat cct gaa atg cgc ccc gtc ttc gtc aac atc ggt gag Leu Leu Leu Asp Pro Glu Met Arg Pro Val Phe Val Asn Ile Gly Glu 420 425 430	1296
cgc acc aac gtc acc ggc tcg gcc aag ttc aag aag ctg atc gtc gaa Arg Thr Asn Val Thr Gly Ser Ala Lys Phe Lys Lys Leu Ile Val Glu 435 440 445	1344
ggg aac tat ccc gag gcg ctg tcg gtc gcg cgc cag cag gtc gag gcc Gly Asn Tyr Pro Glu Ala Leu Ser Val Ala Arg Gln Gln Val Glu Ala 450 455 460	1392
ggg gcc cag gtc atc gac gtg aac atg gac gag ggt ctg ctg gac agc Gly Ala Gln Val Ile Asp Val Asn Met Asp Glu Gly Leu Leu Asp Ser 465 470 475 480	1440
cag cag gcc atg gtc acc ttc ctg aat ctg atg gcg gcc gag ccc gac Gln Gln Ala Met Val Thr Phe Leu Asn Leu Met Ala Ala Glu Pro Asp 485 490 495	1488
atc gcg cgc gtg ccg gtg atg atc gac agc tcc aag tgg gag gtg atc Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys Trp Glu Val Ile 500 505 510	1536
gag gcg ggc ctg aag tgc gta caa ggc aag gcg atc gtc aac tcg atc Glu Ala Gly Leu Lys Cys Val Gln Gly Lys Ala Ile Val Asn Ser Ile 515 520 525	1584
agc ctg aag gaa ggc gag gaa aag ttc ctc gaa cag gcc acg ctc tgc Ser Leu Lys Glu Gly Glu Glu Lys Phe Leu Glu Gln Ala Thr Leu Cys 530 535 540	1632
ctg cgc tat ggc gca gcc gtg gtg gtc atg gcc ttc gac gag gtt ggc Leu Arg Tyr Gly Ala Ala Val Val Val Met Ala Phe Asp Glu Val Gly 545 550 555 560	1680
cag gcc gac acc gaa aag cgc aag gtc gag atc tgt acg cgg gcc tac Gln Ala Asp Thr Glu Lys Arg Lys Val Glu Ile Cys Thr Arg Ala Tyr 565 570 575	1728
aac acg ctc gtg gac aag gtc ggc ttc ccg ccc gag gac atc atc ttc Asn Thr Leu Val Asp Lys Val Gly Phe Pro Pro Glu Asp Ile Ile Phe	1776

173

580	585	590	
gac ccc aac atc ttc gcc gtg gcg acg ggg atc gag gag cac gac aac Asp Pro Asn Ile Phe Ala Val Ala Thr Gly Ile Glu Glu His Asp Asn 595 600 605			1824
tac gcc gtc gac ttc atc gag gcc acg cgg cgc atc aag cag atg ttg Tyr Ala Val Asp Phe Ile Glu Ala Thr Arg Arg Ile Lys Gln Met Leu 610 615 620			1872
ccc tat gcg cgg gtg tcg ggc ggg gtg tcg aac gtc tcg ttc agc ttc Pro Tyr Ala Arg Val Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe 625 630 635 640			1920
cgg ggc aat gag ccg gtg cgc cgg gcg atc cac tcg gtg ttc ctg tac Arg Gly Asn Glu Pro Val Arg Arg Ala Ile His Ser Val Phe Leu Tyr 645 650 655			1968
cac gcc atc aac gcc ggc atg gac atg ggc atc gtc aac gcc ggc gac His Ala Ile Asn Ala Gly Met Asp Met Gly Ile Val Asn Ala Gly Asp 660 665 670			2016
ctg ccg gtc tat gac gac atc gat ccg gcc ctg cgc gag gcc gtc gag Leu Pro Val Tyr Asp Asp Ile Asp Pro Ala Leu Arg Glu Ala Val Glu 675 680 685			2064
gac gtg atc ctc aac ccg ccg cag cgc gat ccg gtg atg acc aac acc Asp Val Ile Leu Asn Arg Pro Gln Arg Asp Pro Val Met Thr Asn Thr 690 695 700			2112
gag cgc ctg gtc gag atg gcc ccg cgc tat aag ggc gag aag ggg cag Glu Arg Leu Val Glu Met Ala Pro Arg Tyr Lys Gly Glu Lys Gly Gln 705 710 715 720			2160
cag cag gtc gcc aac ctg gag tgg cga aag ggc acg gtg aac gag cgc Gln Gln Val Ala Asn Leu Glu Trp Arg Lys Gly Thr Val Asn Glu Arg 725 730 735			2208
ctg acc cat gct ctc gtt cac ggc atc acc gag ttc atc gag cag gac Leu Thr His Ala Leu Val His Gly Ile Thr Glu Phe Ile Glu Gln Asp 740 745 750			2256
acc gag gag gcg cgc ctg gcc gcc gag cgc ccc ttg cac gtg att gaa Thr Glu Glu Ala Arg Leu Ala Ala Glu Arg Pro Leu His Val Ile Glu 755 760 765			2304
ggc ccg ctg atg gac ggc atg aac gtc gtc ggc gac ctg ttc ggc gcg Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala 770 775 780			2352
ggc aag atg ttc ctg ccc cag gtg gtg aag tcg gcc cgc gtg atg aag Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys 785 790 795 800			2400
cag gcc gtc gcc tgg ctg atg ccg ttc atg gag gcc gag aag gaa ggc Gln Ala Val Ala Trp Leu Met Pro Phe Met Glu Ala Glu Lys Glu Gly 805 810 815			2448
cag gag cgc aag gcc gcc ggc aag gtg ctg atg gcc acc gtc aag ggc Gln Glu Arg Lys Ala Ala Gly Lys Val Leu Met Ala Thr Val Lys Gly 820 825 830			2496

174

gac gtc cac gac atc ggt aag aac atc gtc ggc gtc gtg ctg cag tgt	2544
Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln Cys	
835 840 845	
aac aac tac gag gtc gtg gac ctg ggt gtc atg gtg ccc gcc gac cgc	2592
Asn Asn Tyr Glu Val Val Asp Leu Gly Val Met Val Pro Ala Asp Arg	
850 855 860	
atc ctg gac gaa gcc aag aag cac aag gtc gac atg atc ggc ctg tcg	2640
Ile Leu Asp Glu Ala Lys Lys His Lys Val Asp Met Ile Gly Leu Ser	
865 870 875 880	
ggc ctg atc acc ccc tcg ctg gac gag atg gtg ttc gtg gcc gcc gag	2688
Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Phe Val Ala Ala Glu	
885 890 895	
atg gag cgc cag ggc ttt gat atc ccg ctg ctg atc ggc ggc gcc acc	2736
Met Glu Arg Gln Gly Phe Asp Ile Pro Leu Leu Ile Gly Gly Ala Thr	
900 905 910	
acc agc cgc acc cac acc gcg gtg aag atc gag ccg gcc tat cgc cgg	2784
Thr Ser Arg Thr His Thr Ala Val Lys Ile Glu Pro Ala Tyr Arg Arg	
915 920 925	
ggt ccg acg acc tat gtc gtc gac gcc agc cgc gcc gtg ggc gtg gtc	2832
Gly Pro Thr Thr Tyr Val Val Asp Ala Ser Arg Ala Val Gly Val Val	
930 935 940	
tcg ggc ctg ctg tcg gaa ggc gag cgt gac cgg atc atc gcc gag acc	2880
Ser Gly Leu Leu Ser Glu Gly Glu Arg Asp Arg Ile Ile Ala Glu Thr	
945 950 955 960	
cgc gcc gag tat gtg aag gtc cgc gag caa tac gcg cgc ggc cag acc	2928
Arg Ala Glu Tyr Val Lys Val Arg Glu Gln Tyr Ala Arg Gly Gln Thr	
965 970 975	
acc aag gcc cgc gcc tcg atc cag gag gcc cgc aag cgc gcc ttc gcc	2976
Thr Lys Ala Arg Ala Ser Ile Gln Glu Ala Arg Lys Arg Ala Phe Ala	
980 985 990	
att gac tgg aag ggc tat gcg ccg ccc aag ccc gcc ttc atc ggc acg	3024
Ile Asp Trp Lys Gly Tyr Ala Pro Pro Lys Pro Ala Phe Ile Gly Thr	
995 1000 1005	
cgg gtg ttc gag ccg tcg ctg gcc gag ctg gtc ccg ttc atc gac tgg	3072
Arg Val Phe Glu Pro Ser Leu Ala Glu Leu Val Pro Phe Ile Asp Trp	
1010 1015 1020	
tcg ccg ttc ttc gcc agc tgg gag ctg atc ggc cgc ttc ccg cag atc	3120
Ser Pro Phe Phe Ala Ser Trp Glu Leu Ile Gly Arg Phe Pro Gln Ile	
1025 1030 1035 1040	
ctg gag gac gac gtg gtc ggc cag gcc gcc acc gac ctc tac cgc gac	3168
Leu Glu Asp Asp Val Val Gly Gln Ala Ala Thr Asp Leu Tyr Arg Asp	
1045 1050 1055	
gcc cgc gcc atg ctg gac aag gtg gtc gag gaa aag tgg ttc ggg gcc	3216
Ala Arg Ala Met Leu Asp Lys Val Val Glu Lys Trp Phe Gly Ala	
1060 1065 1070	
aag ggc gtg atc ggc ttc tgg ccg gcc cag gcc cag ggc gac gac atc	3264
Lys Gly Val Ile Gly Phe Trp Pro Ala Gln Ala Gln Gly Asp Asp Ile	

175

1075	1080	1085	
gtg ctc tat acc gac gag acc cgc gtg gcc gag ttc tcg cgc ctg cac Val Leu Tyr Thr Asp Glu Thr Arg Val Ala Glu Phe Ser Arg Leu His 1090 1095 1100			3312
acc ctt cgc cag cag atg gac aag ggc gcc gac aag agc ggc gag gcc Thr Leu Arg Gln Gln Met Asp Lys Gly Ala Asp Lys Ser Gly Glu Ala 1105 1110 1115 1120			3360
aag gcc aat gtc gcc ctg tcg gac ttc gtc gcg ccg atc ggg cag ggg Lys Ala Asn Val Ala Leu Ser Asp Phe Val Ala Pro Ile Gly Gln Gly 1125 1130 1135			3408
gct gac tat gtc ggc ggc ttc gcc gtc acc gca ggc cat ggc gag gac Ala Asp Tyr Val Gly Gly Phe Ala Val Thr Ala Gly His Gly Glu Asp 1140 1145 1150			3456
gag atc gtc gcc aag ttc aag gcg gcc ggc gac gac tac aac gcc atc Glu Ile Val Ala Lys Phe Lys Ala Ala Gly Asp Asp Tyr Asn Ala Ile 1155 1160 1165			3504
atg gcc tcg gcc ctg gcc gac cgc ctg gcc gaa gcc ttc gcc gag tgg Met Ala Ser Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Trp 1170 1175 1180			3552
ctg cac tac aaa gcc cgt gtc gag ctg tgg ggc tac gcc gcc gac gag Leu His Tyr Lys Ala Arg Val Glu Leu Trp Gly Tyr Ala Ala Asp Glu 1185 1190 1195 1200			3600
gac gcc gac gtc gag cgc ctg atc gcc gaa aag tac cag ggc atc cgc Asp Ala Asp Val Glu Arg Leu Ile Ala Glu Lys Tyr Gln Gly Ile Arg 1205 1210 1215			3648
ccc gcg ccc ggc tat ccg gcc cag ccc gac cac acc gag aaa ggt acg Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr Glu Lys Gly Thr 1220 1225 1230			3696
ctg ttc aag ctg ctc gac gcc gag gcg gcc acc ggt ctg cag ctg acc Leu Phe Lys Leu Leu Asp Ala Glu Ala Ala Thr Gly Leu Gln Leu Thr 1235 1240 1245			3744
gag agc tac gcc atg acc cct ggc gcg gcg gtc tcc ggc ctg ttc ttc Glu Ser Tyr Ala Met Thr Pro Gly Ala Ala Val Ser Gly Leu Phe Phe 1250 1255 1260			3792
agc cac cgc cag gcg cac tat ttc ggg gtc ggc aag atc gac gcc gac Ser His Arg Gln Ala His Tyr Phe Gly Val Gly Lys Ile Asp Ala Asp 1265 1270 1275 1280			3840
cag gtc gag gac tac gcc cgc cgc aag ggc tgg gat atg gag acg gcc Gln Val Glu Asp Tyr Ala Arg Arg Lys Gly Trp Asp Met Glu Thr Ala 1285 1290 1295			3888
gag cgc tgg ctg tcg ccg atc ctg aac tac gat ccg cta gcg cgg gcg Glu Arg Trp Leu Ser Pro Ile Leu Asn Tyr Asp Pro Leu Ala Arg Ala 1300 1305 1310			3936
cgc ggg gcg gcg gct tag Arg Gly Ala Ala Ala 1315			3954

176

<210> 40  
 <211> 1317  
 <212> PRT  
 <213> *Caulobacter\_crescentus*

<220>  
 <221> unsure  
 <222> 359 .. 359  
 <223> All occurrences of Xaa indicate any amino acid

<220>  
 <221> unsure  
 <222> 416 .. 416  
 <223> All occurrences of Xaa indicate any amino acid

<400> 40  
 Met Thr Asp Leu Ser Ile Arg Ala Asn Arg Val Ala Ala Leu Lys Ala  
   1                  5                  10                  15  
 Ala Ala Lys Glu Arg Ile Leu Ile Leu Asp Gly Ser Trp Gly Val Met  
           20                  25                  30  
 Phe Gln Lys Lys Gly Leu Thr Glu Ala Asp Tyr Arg Ala Glu Arg Phe  
       35                  40                  45  
 Ala Ala Tyr Asn Gly Gln Met Lys Gly Asn Asn Asp Ile Leu Cys Leu  
       50                  55                  60  
 Thr Arg Pro Asp Leu Val Ala Glu Leu His Asp Ala Tyr Phe Ser Ala  
       65                  70                  75                  80  
 Gly Ala Asp Ile Ser Glu Thr Asn Thr Phe Ser Gly Thr Thr Ile Ala  
           85                  90                  95  
 Gln Ala Asp Tyr His Leu Gly Glu Gln Asp Val Trp Asp Ile Asn Leu  
       100                  105                  110  
 Glu Gly Ala Lys Ile Gly Arg Ser Val Ala Asp Arg Trp Asn Ala Gln  
       115                  120                  125  
 Asn Pro Asp Arg Pro Lys Phe Ile Ala Gly Ser Met Gly Pro Leu Asn  
       130                  135                  140  
 Val Met Leu Ser Met Ser Ser Asp Val Asn Asp Pro Gly Ala Arg Lys  
       145                  150                  155                  160  
 Val Thr Phe Asp Gln Val Tyr Glu Ala Tyr Arg Gln Gln Val Asp Ala  
           165                  170                  175  
 Leu Tyr Gln Gly Gly Val Asp Leu Phe Leu Ile Glu Thr Ile Thr Asp  
       180                  185                  190  
 Thr Leu Asn Cys Lys Ala Ala Ile Lys Ala Ile Leu Asp Trp Arg Asp  
       195                  200                  205  
 Glu Gly His Glu Glu Leu Pro Ile Trp Ile Ser Gly Thr Ile Thr Asp  
       210                  215                  220  
 Arg Ser Gly Arg Thr Leu Ser Gly Gln Thr Ala Glu Ala Phe Trp Asn  
       225                  230                  235                  240

177

Ser Val Lys His Ala Lys Pro Phe Ala Val Gly Phe Asn Cys Ala Leu  
 245 250 255  
 Gly Ala Asp Leu Met Arg Pro His Ile Ala Glu Met Ala Arg Ile Ala  
 260 265 270  
 Asp Thr Leu Val Ala Ala Tyr Pro Asn Ala Gly Leu Pro Asn Ala Met  
 275 280 285  
 Gly Gln Tyr Asp Glu Glu Pro His Glu Thr Gly His Ala Leu His Glu  
 290 295 300  
 Trp Ala Lys Asp Gly Leu Val Asn Ile Leu Gly Gly Cys Cys Gly Thr  
 305 310 315 320  
 Thr Pro Asp His Ile Arg His Val Ala Asp Glu Val Arg Gly Val Thr  
 325 330 335  
 Pro Arg Gln Ile Pro Glu Arg Pro Lys Ala Met Arg Leu Ala Gly Leu  
 340 345 350  
 Glu Pro Phe Glu Leu Ala Xaa Trp Leu Arg Pro Gln Ile Pro Phe Ser  
 355 360 365  
 Pro Cys Gly Arg Arg Cys Arg Arg Arg Arg Arg Met Arg Gly Leu Ala  
 370 375 380  
 Gly Pro Ser Thr Ala Val Ser Arg Arg Arg Arg Ser Ser Thr Pro His  
 385 390 395 400  
 Pro Thr Arg Cys Ala Gly His Leu Leu Pro Gln Gly Glu Lys Gly Xaa  
 405 410 415  
 Leu Leu Leu Asp Pro Glu Met Arg Pro Val Phe Val Asn Ile Gly Glu  
 420 425 430  
 Arg Thr Asn Val Thr Gly Ser Ala Lys Phe Lys Lys Leu Ile Val Glu  
 435 440 445  
 Gly Asn Tyr Pro Glu Ala Leu Ser Val Ala Arg Gln Gln Val Glu Ala  
 450 455 460  
 Gly Ala Gln Val Ile Asp Val Asn Met Asp Glu Gly Leu Leu Asp Ser  
 465 470 475 480  
 Gln Gln Ala Met Val Thr Phe Leu Asn Leu Met Ala Ala Glu Pro Asp  
 485 490 495  
 Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys Trp Glu Val Ile  
 500 505 510  
 Glu Ala Gly Leu Lys Cys Val Gln Gly Lys Ala Ile Val Asn Ser Ile  
 515 520 525  
 Ser Leu Lys Glu Gly Glu Glu Lys Phe Leu Glu Gln Ala Thr Leu Cys  
 530 535 540  
 Leu Arg Tyr Gly Ala Ala Val Val Val Met Ala Phe Asp Glu Val Gly  
 545 550 555 560  
 Gln Ala Asp Thr Glu Lys Arg Lys Val Glu Ile Cys Thr Arg Ala Tyr  
 565 570 575

178

Asn Thr Leu Val Asp Lys Val Gly Phe Pro Pro Glu Asp Ile Ile Phe  
 580 585 590  
 Asp Pro Asn Ile Phe Ala Val Ala Thr Gly Ile Glu Glu His Asp Asn  
 595 600 605  
 Tyr Ala Val Asp Phe Ile Glu Ala Thr Arg Arg Ile Lys Gln Met Leu  
 610 615 620  
 Pro Tyr Ala Arg Val Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe  
 625 630 635 640  
 Arg Gly Asn Glu Pro Val Arg Arg Ala Ile His Ser Val Phe Leu Tyr  
 645 650 655  
 His Ala Ile Asn Ala Gly Met Asp Met Gly Ile Val Asn Ala Gly Asp  
 660 665 670  
 Leu Pro Val Tyr Asp Asp Ile Asp Pro Ala Leu Arg Glu Ala Val Glu  
 675 680 685  
 Asp Val Ile Leu Asn Arg Pro Gln Arg Asp Pro Val Met Thr Asn Thr  
 690 695 700  
 Glu Arg Leu Val Glu Met Ala Pro Arg Tyr Lys Gly Glu Lys Gly Gln  
 705 710 715 720  
 Gln Gln Val Ala Asn Leu Glu Trp Arg Lys Gly Thr Val Asn Glu Arg  
 725 730 735  
 Leu Thr His Ala Leu Val His Gly Ile Thr Glu Phe Ile Glu Gln Asp  
 740 745 750  
 Thr Glu Glu Ala Arg Leu Ala Ala Glu Arg Pro Leu His Val Ile Glu  
 755 760 765  
 Gly Pro Leu Met Asp Gly Met Asn Val Val Gly Asp Leu Phe Gly Ala  
 770 775 780  
 Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met Lys  
 785 790 795 800  
 Gln Ala Val Ala Trp Leu Met Pro Phe Met Glu Ala Glu Lys Glu Gly  
 805 810 815  
 Gln Glu Arg Lys Ala Ala Gly Lys Val Leu Met Ala Thr Val Lys Gly  
 820 825 830  
 Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gln Cys  
 835 840 845  
 Asn Asn Tyr Glu Val Val Asp Leu Gly Val Met Val Pro Ala Asp Arg  
 850 855 860  
 Ile Leu Asp Glu Ala Lys Lys His Lys Val Asp Met Ile Gly Leu Ser  
 865 870 875 880  
 Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val Phe Val Ala Ala Glu  
 885 890 895  
 Met Glu Arg Gln Gly Phe Asp Ile Pro Leu Leu Ile Gly Gly Ala Thr



179

900	905	910
Thr Ser Arg Thr His Thr Ala Val Lys Ile Glu Pro Ala Tyr Arg Arg 915 920 925		
Gly Pro Thr Thr Tyr Val Val Asp Ala Ser Arg Ala Val Gly Val Val 930 935 940		
Ser Gly Leu Leu Ser Glu Gly Glu Arg Asp Arg Ile Ile Ala Glu Thr 945 950 955 960		
Arg Ala Glu Tyr Val Lys Val Arg Glu Gln Tyr Ala Arg Gly Gln Thr 965 970 975		
Thr Lys Ala Arg Ala Ser Ile Gln Glu Ala Arg Lys Arg Ala Phe Ala 980 985 990		
Ile Asp Trp Lys Gly Tyr Ala Pro Pro Lys Pro Ala Phe Ile Gly Thr 995 1000 1005		
Arg Val Phe Glu Pro Ser Leu Ala Glu Leu Val Pro Phe Ile Asp Trp 1010 1015 1020		
Ser Pro Phe Phe Ala Ser Trp Glu Leu Ile Gly Arg Phe Pro Gln Ile 1025 1030 1035 1040		
Leu Glu Asp Asp Val Val Gly Gln Ala Ala Thr Asp Leu Tyr Arg Asp 1045 1050 1055		
Ala Arg Ala Met Leu Asp Lys Val Val Glu Glu Lys Trp Phe Gly Ala 1060 1065 1070		
Lys Gly Val Ile Gly Phe Trp Pro Ala Gln Ala Gln Gly Asp Asp Ile 1075 1080 1085		
Val Leu Tyr Thr Asp Glu Thr Arg Val Ala Glu Phe Ser Arg Leu His 1090 1095 1100		
Thr Leu Arg Gln Gln Met Asp Lys Gly Ala Asp Lys Ser Gly Glu Ala 1105 1110 1115 1120		
Lys Ala Asn Val Ala Leu Ser Asp Phe Val Ala Pro Ile Gly Gln Gly 1125 1130 1135		
Ala Asp Tyr Val Gly Gly Phe Ala Val Thr Ala Gly His Gly Glu Asp 1140 1145 1150		
Glu Ile Val Ala Lys Phe Lys Ala Ala Gly Asp Asp Tyr Asn Ala Ile 1155 1160 1165		
Met Ala Ser Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu Trp 1170 1175 1180		
Leu His Tyr Lys Ala Arg Val Glu Leu Trp Gly Tyr Ala Ala Asp Glu 1185 1190 1195 1200		
Asp Ala Asp Val Glu Arg Leu Ile Ala Glu Lys Tyr Gln Gly Ile Arg 1205 1210 1215		
Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr Glu Lys Gly Thr 1220 1225 1230		

180

Leu Phe Lys Leu Leu Asp Ala Glu Ala Ala Thr Gly Leu Gln Leu Thr  
 1235 1240 1245

Glu Ser Tyr Ala Met Thr Pro Gly Ala Ala Val Ser Gly Leu Phe Phe  
 1250 1255 1260

Ser His Arg Gln Ala His Tyr Phe Gly Val Gly Lys Ile Asp Ala Asp  
 1265 1270 1275 1280

Gln Val Glu Asp Tyr Ala Arg Arg Lys Gly Trp Asp Met Glu Thr Ala  
 1285 1290 1295

Glu Arg Trp Leu Ser Pro Ile Leu Asn Tyr Asp Pro Leu Ala Arg Ala  
 1300 1305 1310

Arg Gly Ala Ala Ala  
 1315

<210> 41  
 <211> 3759  
 <212> DNA  
 <213> Rhodobacter capsulatus

<220>  
 <221> CDS  
 <222> (1) .. (3756)  
 <223> RRC01731

<400> 41  
 atg ctg acc cag acc ctg ccc cga tct gcg gcc ttt gcc gca att gag 48  
 Met Leu Thr Gln Thr Leu Pro Arg Ser Ala Ala Phe Ala Ala Ile Glu  
 1 5 10 15  
 gcg ctt tcg cgc cag cgg atc ttg atc ctt gac ggg gcg atg ggc acg 96  
 Ala Leu Ser Arg Gln Arg Ile Leu Ile Leu Asp Gly Ala Met Gly Thr  
 20 25 30  
 cag atc cag cag ctt ggc ctg agc gag gac gat ttt ctg ggc cac ggc 144  
 Gln Ile Gln Gln Leu Gly Leu Ser Glu Asp Asp Phe Leu Gly His Gly  
 35 40 45  
 tcg ggc tgc gcc tgc cgc cat gcc acc gat cat ccg caa aag ggc aac 192  
 Ser Gly Cys Ala Cys Arg His Ala Thr Asp His Pro Gln Lys Gly Asn  
 50 55 60  
 aac gac ctg ctg gtg ctg acc cag ccg caa gcg atc gag gag atc cat 240  
 Asn Asp Leu Leu Val Leu Thr Gln Pro Gln Ala Ile Glu Glu Ile His  
 65 70 75 80  
 ttc cgc tat gcg atg gcg ggg gcg gat atc gtc gag acg aac acc ttt 288  
 Phe Arg Tyr Ala Met Ala Gly Ala Asp Ile Val Glu Thr Asn Thr Phe  
 85 90 95  
 tcg gcc acc acc atc gcg cag gcc gat tac ggg ctg gaa agc gcg gtg 336  
 Ser Ala Thr Thr Ile Ala Gln Ala Asp Tyr Gly Leu Glu Ser Ala Val  
 100 105 110  
 ttc gac ctg aac gcc gcg ggg gcg cgg gtg gcg cgg gcg gcg atg gac 384  
 Phe Asp Leu Asn Ala Ala Gly Ala Arg Val Ala Arg Ala Ala Met Asp  
 115 120 125

181

cgc gcc gag gcc acc gac gga cgg cgc cgc ttc gtt gcg ggg gcg gtg Arg Ala Glu Ala Thr Asp Gly Arg Arg Arg Phe Val Ala Gly Ala Val 130 135 140	432
ggg ccg acg aac cgc acc gcc tcg ctc tcg ccc gat gtg aac gac ccg Gly Pro Thr Asn Arg Thr Ala Ser Leu Ser Pro Asp Val Asn Asp Pro 145 150 155 160	480
ggc ttt cgc gcc gtc acc ttc gac gat ctg cgc acg gcc tat ggc cag Gly Phe Arg Ala Val Thr Phe Asp Asp Leu Arg Thr Ala Tyr Gly Gln 165 170 175	528
cag gtg cgc ggt ctg atc gcg ggg ggc gcc gat atc ctg ctg atc gag Gln Val Arg Gly Leu Ile Ala Gly Gly Ala Asp Ile Leu Leu Ile Glu 180 185 190	576
acg atc ttt gac acg ctg aac gcc aag gcg gcg att ttc gcc tgt ttc Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Cys Phe 195 200 205	624
gaa gcc ttt gcc gaa cgg ggc gag cgg ctg ccg gtg atg att tcc ggc Glu Ala Phe Ala Glu Arg Gly Glu Arg Leu Pro Val Met Ile Ser Gly 210 215 220	672
acg atc acc gat gcc tcg ggg cgc aca ttg tcg ggg cag acg ccg acc Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Pro Thr 225 230 235 240	720
gcg ttc tgg cat tcg gtg gct cat gcc cgg ccc ttt acc gtg ggg ctg Ala Phe Trp His Ser Val Ala His Ala Arg Pro Phe Thr Val Gly Leu 245 250 255	768
aac tgc gcg ctg ggc gcc agt gcg atg cgt ccg cat ctg gcg gaa ctg Asn Cys Ala Leu Gly Ala Ser Ala Met Arg Pro His Leu Ala Glu Leu 260 265 270	816
gcg ggc gtc gcc ccc tgc gcg atc tgc gcc tat ccc aat gcc ggg ctg Ala Gly Val Ala Pro Cys Ala Ile Cys Ala Tyr Pro Asn Ala Gly Leu 275 280 285	864
ccc aat gcc ttt ggc caa tat gac gaa acc ccc gac cgg acc gcc gcg Pro Asn Ala Phe Gly Gln Tyr Asp Glu Thr Pro Asp Arg Thr Ala Ala 290 295 300	912
cag gtg gcc gaa ttt gcc cgc gaa ggg ctg gtc aat gtc gtg ggc ggt Gln Val Ala Glu Phe Ala Arg Glu Gly Leu Val Asn Val Val Gly Gly 305 310 315 320	960
tgc tgc ggc acc acc ccc gat cac atc cgc gcc atc gcg gaa gcc gtg Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Ile Ala Glu Ala Val 325 330 335	1008
aaa cct ttc ccg ccg agg gcc ctg cca agc cgt tat ctg cgc ctt tcg Lys Pro Phe Pro Pro Arg Ala Leu Pro Ser Arg Tyr Leu Arg Leu Ser 340 345 350	1056
ggg ctt gag ccc ttt acc ctg acg ccc gac att ccc ttc gtg aac atc Gly Leu Glu Pro Phe Thr Leu Thr Pro Asp Ile Pro Phe Val Asn Ile 355 360 365	1104
ggc gag cgc acg aat gtc acc ggc tcg gcc cgg ttc cgc aag atg atc Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Arg Phe Arg Lys Met Ile	1152

182

370	375	380	
gtc gcc cgc gac tat Val Ala Arg Asp Tyr 385	gcc gcc gcg ctg gat Ala Ala Ala Leu Asp 390	gtc gcc cgc gat cag gtg Val Ala Arg Asp Gln Val 395 400	1200
gaa aac ggc gcg cag atc ctt gac atc aac atg gac gag ggg ctg atc Glu Asn Gly Ala Gln Ile Leu Asp Ile Asn Met Asp Glu Gly Leu Ile 405 410 415			1248
gac agt cag gcg gcg atg gtc gcc ttc ctc aac ctc ttg gcc gcc gag Asp Ser Gln Ala Ala Met Val Ala Phe Leu Asn Leu Leu Ala Ala Glu 420 425 430			1296
ccc gac att gcc cgg gtg ccg gtg atg atc gac agc tcg aaa tgg gag Pro Asp Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys Trp Glu 435 440 445			1344
gtg atc gag gcc ggg ctg aaa tgc gtg cag ggc aag ccc gtc gtc aat Val Ile Glu Ala Gly Leu Lys Cys Val Gln Gly Lys Pro Val Val Asn 450 455 460			1392
tcg atc agc ctg aag gag ggc gag gag atc ttc cgc cat cac gcg gcg Ser Ile Ser Leu Lys Glu Gly Glu Glu Ile Phe Arg His His Ala Ala 465 470 475 480			1440
ctg tgt ctg gcc tat ggc gcg gcg gtc gtc gtg atg gcc ttt gac gaa Leu Cys Leu Ala Tyr Gly Ala Ala Val Val Val Met Ala Phe Asp Glu 485 490 495			1488
gag ggg cag gcc gac agt ttc gcc cga aag acc agc atc tgc gcc cgc Glu Gly Gln Ala Asp Ser Phe Ala Arg Lys Thr Ser Ile Cys Ala Arg 500 505 510			1536
gcc tat cgc att ctg gtc gag gag atc ggc ttt ccg ccc gaa gac atc Ala Tyr Arg Ile Leu Val Glu Glu Ile Gly Phe Pro Pro Glu Asp Ile 515 520 525			1584
atc ttt gac ccg aac gtc ttt gcc gtc gcc acg ggc atc gaa gaa cac Ile Phe Asp Pro Asn Val Phe Ala Val Ala Thr Gly Ile Glu Glu His 530 535 540			1632
gac aat tac ggc gtt gat ttc atc gag gcc gct cgc tgg atc cgg gcc Asp Asn Tyr Gly Val Asp Phe Ile Glu Ala Ala Arg Trp Ile Arg Ala 545 550 555 560			1680
aac ctg ccg cat gcc cat gtc tcg ggc ggg gtg tcg aac ctg tcc ttc Asn Leu Pro His Ala His Val Ser Gly Gly Val Ser Asn Leu Ser Phe 565 570 575			1728
agc ttt cgc ggc aac gaa ccc gtg cgc gcg gcg atg cat gcg gtg ttt Ser Phe Arg Gly Asn Glu Pro Val Arg Ala Ala Met His Ala Val Phe 580 585 590			1776
ctt tac cac gcc atc cgc gcc ggg atg gat atg ggg atc gtc aat gcc Leu Tyr His Ala Ile Arg Ala Gly Met Asp Met Gly Ile Val Asn Ala 595 600 605			1824
ggg cag ctg gtg gtc tat gac cag atc gac ccc gag ctg cgc cag gcc Gly Gln Leu Val Val Tyr Asp Gln Ile Asp Pro Glu Leu Arg Gln Ala 610 615 620			1872

183

tgc gag gat gtg gtg ctc aac cgc cag ccc aaa tcg ggc ggc acc gcg Cys Glu Asp Val Val Leu Asn Arg Gln Pro Lys Ser Gly Gly Thr Ala 625 630 635 640	1920
acc gag cgg atg ctg gag gtg gcc gag cgc ttc cgc ggc ggc gcg cgc Thr Glu Arg Met Leu Glu Val Ala Glu Arg Phe Arg Gly Gly Ala Arg 645 650 655	1968
gag gaa aag acc cgc gat ctg gcc tgg cgc gac tgg ccg gtg gaa aag Glu Glu Lys Thr Arg Asp Leu Ala Trp Arg Asp Trp Pro Val Glu Lys 660 665 670	2016
cgg ctc gaa cat gcg ctg gtc aat ggc atc acc gaa ttc atc gag gcc Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Glu Phe Ile Glu Ala 675 680 685	2064
gat acc gaa gcc gca agg ctt ctg gcc gaa cgc ccg ctg cat gtg atc Asp Thr Glu Ala Ala Arg Leu Leu Ala Glu Arg Pro Leu His Val Ile 690 695 700	2112
gaa ggg ccg ctg atg gcg ggg atg aat gtc gtc ggt gat ctg ttc ggc Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp Leu Phe Gly 705 710 715 720	2160
gcg ggc aag atg ttc ctg cca cag gtg gtg aaa tcg gcg cgc gtg atg Ala Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met 725 730 735	2208
aaa cag gcc gtc gcc gtt ctg ctg ccc tac atg gat gcc gaa aag gcc Lys Gln Ala Val Ala Val Leu Leu Pro Tyr Met Asp Ala Glu Lys Ala 740 745 750	2256
gcg cgc ggc ggc gag ggg cgc gaa acc gcg ggc aag atc ctg atg gcc Ala Arg Gly Gly Glu Gly Arg Glu Thr Ala Gly Lys Ile Leu Met Ala 755 760 765	2304
acg gtc aag ggc gat gtg cat gac atc ggc aag aac atc gtc ggc gtc Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val 770 775 780	2352
gtg ctg gcc tgc aac aat tac gac atc gtc gac ctg ggc gtg atg gtg Val Leu Ala Cys Asn Asn Tyr Asp Ile Val Asp Leu Gly Val Met Val 785 790 795 800	2400
ccg ccg caa aag atc ctg gaa gtg gcg cgg gcc gaa aag gtc gat gcg Pro Pro Gln Lys Ile Leu Glu Val Ala Arg Ala Glu Lys Val Asp Ala 805 810 815	2448
atc ggg ctt tcc ggg ctg atc acg cca agc ctg gac gag atg gtg cat Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val His 820 825 830	2496
ctg gcc gcg gaa atg gag cgc gag ggc ttt gac att ccg ctg ctg atc Leu Ala Ala Glu Met Glu Arg Glu Gly Phe Asp Ile Pro Leu Leu Ile 835 840 845	2544
ggc ggg gcg acc acg tcg aaa gtg cat acg gcg gtg aag atc gcc ccc Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Ile Ala Pro 850 855 860	2592
gcc tac agc cgc ggg cag gcg gtt tat gtg ctc gat gcc agc cgg gcc Ala Tyr Ser Arg Gly Gln Ala Val Tyr Val Leu Asp Ala Ser Arg Ala 865 870 875 880	2640

184

865	870	875	880	
gtg ggg gtg gtg ggg gcg ctt ttg agc ccg aac cag aag gtc gat tac	2688			
Val Gly Val Val Gly Ala Leu Leu Ser Pro Asn Gln Lys Val Asp Tyr				
885	890	895		
gcg gcg cag atc cgc gcg gac tat gcg cag atc gcc gcc cgt cat gcc	2736			
Ala Ala Gln Ile Arg Ala Asp Tyr Ala Gln Ile Ala Ala Arg His Ala				
900	905	910		
cgc gac gag gcc gcc aag gtg cgg ctg cct ttg gcc gcg gcc cgg gcc	2784			
Arg Asp Glu Ala Ala Lys Val Arg Leu Pro Leu Ala Ala Ala Arg Ala				
915	920	925		
aat gcg ctg cgg ctc gac tgg tgg ggc tat gcc gtg ccc gcg ccg caa	2832			
Asn Ala Leu Arg Leu Asp Trp Ser Gly Tyr Ala Val Pro Ala Pro Gln				
930	935	940		
ttc ctt ggc ccg cgc gtg atc gac gac tgg gat ctg gcc gaa gtg gcg	2880			
Phe Leu Gly Pro Arg Val Ile Asp Asp Trp Asp Leu Ala Glu Val Ala				
945	950	955	960	
cgg tat atc gac tgg acg ccc ttc ttc cat gcc tgg gaa ttg aag ggg	2928			
Arg Tyr Ile Asp Trp Thr Pro Phe Phe His Ala Trp Glu Leu Lys Gly				
965	970	975		
gtc tat ccg cgg att ctc gat gac gcc gaa aag ggc gaa gcg gcg cgg	2976			
Val Tyr Pro Arg Ile Leu Asp Asp Ala Glu Lys Gly Glu Ala Ala Arg				
980	985	990		
gca ctt ttc gcc gat gcc cag gcg atg ctg gcg cag atc att gcc gaa	3024			
Ala Leu Phe Ala Asp Ala Gln Ala Met Leu Ala Gln Ile Ile Ala Glu				
995	1000	1005		
cgc tgg ttc acc ccg cgc gcc gtg gtg ggg ttc tgg ccc gcg cag gcg	3072			
Arg Trp Phe Thr Pro Arg Ala Val Val Gly Phe Trp Pro Ala Gln Ala				
1010	1015	1020		
gtg ggc gac gat atc cgg ctt tac acc gac gag agc cgg acc gaa gac	3120			
Val Gly Asp Asp Ile Arg Leu Tyr Thr Asp Glu Ser Arg Thr Glu Asp				
1025	1030	1035	1040	
ctc gcc act ttc ttc acc ctg cgc cag cag acc ggc aag cgc gaa ggc	3168			
Leu Ala Thr Phe Phe Thr Leu Arg Gln Gln Thr Gly Lys Arg Glu Gly				
1045	1050	1055		
cgc ccg aat gtg gct ttg gcc gat ttc gtc gcg cct gcg ggc acg gtg	3216			
Arg Pro Asn Val Ala Leu Ala Asp Phe Val Ala Pro Ala Gly Thr Val				
1060	1065	1070		
ccc gat tat ctg ggc ggc ttc gtg gtc acc gcg ggc ccc gag gaa gcc	3264			
Pro Asp Tyr Leu Gly Gly Phe Val Val Thr Ala Gly Pro Glu Glu Ala				
1075	1080	1085		
gag atc gcc gcg cgg ttc gaa gct gcc aat gac cat tat tcc gcg atc	3312			
Glu Ile Ala Ala Arg Phe Glu Ala Ala Asn Asp His Tyr Ser Ala Ile				
1090	1095	1100		
ctg gtc aag gcg ctg gcc gac cgc ttt gcc gaa gcc ctg gcc gag gcc	3360			
Leu Val Lys Ala Leu Ala Asp Arg Phe Ala Glu Ala Leu Ala Glu Ala				
1105	1110	1115	1120	

185

ctg cat cag cgg gtg cgg cgc gac tat tgg ggc tat gcg ccc gaa gaa 3408  
 Leu His Gln Arg Val Arg Arg Asp Tyr Trp Gly Tyr Ala Pro Glu Glu  
                   1125                  1130                  1135

agc ttc gcc ccc gat cag ctg gtg ggc gag ccc tat cgc ggc atc cgc 3456  
 Ser Phe Ala Pro Asp Gln Leu Val Gly Glu Pro Tyr Arg Gly Ile Arg  
                   1140                  1145                  1150

ccg gcg ccc ggc tat ccg gcc cag ccc gac cac acg gaa aag ctg acg 3504  
 Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr Glu Lys Leu Thr  
                   1155                  1160                  1165

ctg ttc cgg ctg ctt ggg gcc gag gcc gcg acc ggc gtg cat ctg acc 3552  
 Leu Phe Arg Leu Leu Gly Ala Glu Ala Ala Thr Gly Val His Leu Thr  
                   1170                  1175                  1180

gac agc atg gcg atg tgg ccc ggc tct tcg gtc tcg ggg ctc tat atc 3600  
 Asp Ser Met Ala Met Trp Pro Gly Ser Ser Val Ser Gly Leu Tyr Ile  
                   1185                  1190                  1195                  1200

ggc cat ccg gag gcc tat tat ttc ggt ctg gcc cgg atc gag cag gat 3648  
 Gly His Pro Glu Ala Tyr Tyr Phe Gly Leu Ala Arg Ile Glu Gln Asp  
                   1205                  1210                  1215

cag gcc gcc gat tac gcc gcc cgc aag ggc atg gcc ttg gcc gag gtg 3696  
 Gln Ala Ala Asp Tyr Ala Ala Arg Lys Gly Met Ala Leu Ala Glu Val  
                   1220                  1225                  1230

cag cgc tgg ctg gcc ccg gtg ctg ggg tcg gcc gcg ccc gcc gcc gct 3744  
 Gln Arg Trp Leu Ala Pro Val Leu Gly Ser Ala Ala Pro Ala Ala Ala  
                   1235                  1240                  1245

gcg gtg gcc gcg tga 3759  
 Ala Val Ala Ala  
                   1250

&lt;210&gt; 42

&lt;211&gt; 1252

&lt;212&gt; PRT

&lt;213&gt; Rhodobacter capsulatus

&lt;400&gt; 42

Met Leu Thr Gln Thr Leu Pro Arg Ser Ala Ala Phe Ala Ala Ile Glu  
   1                  5                  10                  15

Ala Leu Ser Arg Gln Arg Ile Leu Ile Leu Asp Gly Ala Met Gly Thr  
           20                  25                  30

Gln Ile Gln Gln Leu Gly Leu Ser Glu Asp Asp Phe Leu Gly His Gly  
           35                  40                  45

Ser Gly Cys Ala Cys Arg His Ala Thr Asp His Pro Gln Lys Gly Asn  
           50                  55                  60

Asn Asp Leu Leu Val Leu Thr Gln Pro Gln Ala Ile Glu Glu Ile His  
           65                  70                  75                  80

Phe Arg Tyr Ala Met Ala Gly Ala Asp Ile Val Glu Thr Asn Thr Phe  
                   85                  90                  95

Ser Ala Thr Thr Ile Ala Gln Ala Asp Tyr Gly Leu Glu Ser Ala Val

186

100	105	110
Phe Asp Leu Asn Ala Ala Gly Ala Arg Val Ala Arg Ala Ala Met Asp 115 120 125		
Arg Ala Glu Ala Thr Asp Gly Arg Arg Arg Phe Val Ala Gly Ala Val 130 135 140		
Gly Pro Thr Asn Arg Thr Ala Ser Leu Ser Pro Asp Val Asn Asp Pro 145 150 155 160		
Gly Phe Arg Ala Val Thr Phe Asp Asp Leu Arg Thr Ala Tyr Gly Gln 165 170 175		
Gln Val Arg Gly Leu Ile Ala Gly Gly Ala Asp Ile Leu Leu Ile Glu 180 185 190		
Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile Phe Ala Cys Phe 195 200 205		
Glu Ala Phe Ala Glu Arg Gly Glu Arg Leu Pro Val Met Ile Ser Gly 210 215 220		
Thr Ile Thr Asp Ala Ser Gly Arg Thr Leu Ser Gly Gln Thr Pro Thr 225 230 235 240		
Ala Phe Trp His Ser Val Ala His Ala Arg Pro Phe Thr Val Gly Leu 245 250 255		
Asn Cys Ala Leu Gly Ala Ser Ala Met Arg Pro His Leu Ala Glu Leu 260 265 270		
Ala Gly Val Ala Pro Cys Ala Ile Cys Ala Tyr Pro Asn Ala Gly Leu 275 280 285		
Pro Asn Ala Phe Gly Gln Tyr Asp Glu Thr Pro Asp Arg Thr Ala Ala 290 295 300		
Gln Val Ala Glu Phe Ala Arg Glu Gly Leu Val Asn Val Val Gly Gly 305 310 315 320		
Cys Cys Gly Thr Thr Pro Asp His Ile Arg Ala Ile Ala Glu Ala Val 325 330 335		
Lys Pro Phe Pro Pro Arg Ala Leu Pro Ser Arg Tyr Leu Arg Leu Ser 340 345 350		
Gly Leu Glu Pro Phe Thr Leu Thr Pro Asp Ile Pro Phe Val Asn Ile 355 360 365		
Gly Glu Arg Thr Asn Val Thr Gly Ser Ala Arg Phe Arg Lys Met Ile 370 375 380		
Val Ala Arg Asp Tyr Ala Ala Ala Leu Asp Val Ala Arg Asp Gln Val 385 390 395 400		
Glu Asn Gly Ala Gln Ile Leu Asp Ile Asn Met Asp Glu Gly Leu Ile 405 410 415		
Asp Ser Gln Ala Ala Met Val Ala Phe Leu Asn Leu Leu Ala Ala Glu 420 425 430		



187

Pro Asp Ile Ala Arg Val Pro Val Met Ile Asp Ser Ser Lys Trp Glu  
 435 440 445  
 Val Ile Glu Ala Gly Leu Lys Cys Val Gln Gly Lys Pro Val Val Asn  
 450 455 460  
 Ser Ile Ser Leu Lys Glu Gly Glu Glu Ile Phe Arg His His Ala Ala  
 465 470 475 480  
 Leu Cys Leu Ala Tyr Gly Ala Ala Val Val Val Met Ala Phe Asp Glu  
 485 490 495  
 Glu Gly Gln Ala Asp Ser Phe Ala Arg Lys Thr Ser Ile Cys Ala Arg  
 500 505 510  
 Ala Tyr Arg Ile Leu Val Glu Glu Ile Gly Phe Pro Pro Glu Asp Ile  
 515 520 525  
 Ile Phe Asp Pro Asn Val Phe Ala Val Ala Thr Gly Ile Glu Glu His  
 530 535 540  
 Asp Asn Tyr Gly Val Asp Phe Ile Glu Ala Ala Arg Trp Ile Arg Ala  
 545 550 555 560  
 Asn Leu Pro His Ala His Val Ser Gly Gly Val Ser Asn Leu Ser Phe  
 565 570 575  
 Ser Phe Arg Gly Asn Glu Pro Val Arg Ala Ala Met His Ala Val Phe  
 580 585 590  
 Leu Tyr His Ala Ile Arg Ala Gly Met Asp Met Gly Ile Val Asn Ala  
 595 600 605  
 Gly Gln Leu Val Val Tyr Asp Gln Ile Asp Pro Glu Leu Arg Gln Ala  
 610 615 620  
 Cys Glu Asp Val Val Leu Asn Arg Gln Pro Lys Ser Gly Gly Thr Ala  
 625 630 635 640  
 Thr Glu Arg Met Leu Glu Val Ala Glu Arg Phe Arg Gly Gly Ala Arg  
 645 650 655  
 Glu Glu Lys Thr Arg Asp Leu Ala Trp Arg Asp Trp Pro Val Glu Lys  
 660 665 670  
 Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Glu Phe Ile Glu Ala  
 675 680 685  
 Asp Thr Glu Ala Ala Arg Leu Leu Ala Glu Arg Pro Leu His Val Ile  
 690 695 700  
 Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp Leu Phe Gly  
 705 710 715 720  
 Ala Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala Arg Val Met  
 725 730 735  
 Lys Gln Ala Val Ala Val Leu Leu Pro Tyr Met Asp Ala Glu Lys Ala  
 740 745 750  
 Ala Arg Gly Gly Glu Gly Arg Glu Thr Ala Gly Lys Ile Leu Met Ala  
 755 760 765

188

Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Gly Val  
 770 775 780  
 Val Leu Ala Cys Asn Asn Tyr Asp Ile Val Asp Leu Gly Val Met Val  
 785 790 795 800  
 Pro Pro Gln Lys Ile Leu Glu Val Ala Arg Ala Glu Lys Val Asp Ala  
 805 810 815  
 Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met Val His  
 820 825 830  
 Leu Ala Ala Glu Met Glu Arg Glu Gly Phe Asp Ile Pro Leu Leu Ile  
 835 840 845  
 Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Ile Ala Pro  
 850 855 860  
 Ala Tyr Ser Arg Gly Gln Ala Val Tyr Val Leu Asp Ala Ser Arg Ala  
 865 870 875 880  
 Val Gly Val Val Gly Ala Leu Leu Ser Pro Asn Gln Lys Val Asp Tyr  
 885 890 895  
 Ala Ala Gln Ile Arg Ala Asp Tyr Ala Gln Ile Ala Ala Arg His Ala  
 900 905 910  
 Arg Asp Glu Ala Ala Lys Val Arg Leu Pro Leu Ala Ala Ala Arg Ala  
 915 920 925  
 Asn Ala Leu Arg Leu Asp Trp Ser Gly Tyr Ala Val Pro Ala Pro Gln  
 930 935 940  
 Phe Leu Gly Pro Arg Val Ile Asp Asp Trp Asp Leu Ala Glu Val Ala  
 945 950 955 960  
 Arg Tyr Ile Asp Trp Thr Pro Phe Phe His Ala Trp Glu Leu Lys Gly  
 965 970 975  
 Val Tyr Pro Arg Ile Leu Asp Asp Ala Glu Lys Gly Glu Ala Ala Arg  
 980 985 990  
 Ala Leu Phe Ala Asp Ala Gln Ala Met Leu Ala Gln Ile Ile Ala Glu  
 995 1000 1005  
 Arg Trp Phe Thr Pro Arg Ala Val Val Gly Phe Trp Pro Ala Gln Ala  
 1010 1015 1020  
 Val Gly Asp Asp Ile Arg Leu Tyr Thr Asp Glu Ser Arg Thr Glu Asp  
 1025 1030 1035 1040  
 Leu Ala Thr Phe Phe Thr Leu Arg Gln Gln Thr Gly Lys Arg Glu Gly  
 1045 1050 1055  
 Arg Pro Asn Val Ala Leu Ala Asp Phe Val Ala Pro Ala Gly Thr Val  
 1060 1065 1070  
 Pro Asp Tyr Leu Gly Gly Phe Val Val Thr Ala Gly Pro Glu Glu Ala  
 1075 1080 1085  
 Glu Ile Ala Ala Arg Phe Glu Ala Ala Asn Asp His Tyr Ser Ala Ile

189

1090	1095	1100
Leu Val Lys Ala Leu Ala Asp Arg Phe Ala Glu Ala Leu Ala Glu Ala		
1105	1110	1115 1120
Leu His Gln Arg Val Arg Arg Asp Tyr Trp Gly Tyr Ala Pro Glu Glu		
	1125	1130 1135
Ser Phe Ala Pro Asp Gln Leu Val Gly Glu Pro Tyr Arg Gly Ile Arg		
	1140	1145 1150
Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr Glu Lys Leu Thr		
	1155	1160 1165
Leu Phe Arg Leu Leu Gly Ala Glu Ala Ala Thr Gly Val His Leu Thr		
	1170	1175 1180
Asp Ser Met Ala Met Trp Pro Gly Ser Ser Val Ser Gly Leu Tyr Ile		
	1185	1190 1195 1200
Gly His Pro Glu Ala Tyr Tyr Phe Gly Leu Ala Arg Ile Glu Gln Asp		
	1205	1210 1215
Gln Ala Ala Asp Tyr Ala Ala Arg Lys Gly Met Ala Leu Ala Glu Val		
	1220	1225 1230
Gln Arg Trp Leu Ala Pro Val Leu Gly Ser Ala Ala Pro Ala Ala Ala		
	1235	1240 1245
Ala Val Ala Ala		
1250		

<210> 43  
 <211> 3798  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (1) .. (3795)  
 <223> RHS24705

<400> 43

atg tca ccc gcg ctc caa gac ctg tcg caa ccc gaa ggt ctg aag aaa	48
Met Ser Pro Ala Leu Gln Asp Leu Ser Gln Pro Glu Gly Leu Lys Lys	
1 5 10 15	
acc ctg cgg gat gag atc aat gcc att ctg cag aag agg att atg gtg	96
Thr Leu Arg Asp Glu Ile Asn Ala Ile Leu Gln Lys Arg Ile Met Val	
20 25 30	
ctg gat gga ggg atg ggg acc atg atc cag cgg gag aag cta aac gaa	144
Leu Asp Gly Gly Met Gly Thr Met Ile Gln Arg Glu Lys Leu Asn Glu	
35 40 45	
gaa cac ttc cga ggt cag gaa ttt aaa gat cat gcc agg ccg ctg aaa	192
Glu His Phe Arg Gly Gln Glu Phe Lys Asp His Ala Arg Pro Leu Lys	
50 55 60	
ggc aac aat gac att tta agt ata act cag cct gat gtc att tac caa	240
Gly Asn Asn Asp Ile Leu Ser Ile Thr Gln Pro Asp Val Ile Tyr Gln	

65	70	190	75	80	
atc cat aag gaa tac ttg ctg gct ggg gca gat atc att gaa aca aat					288
Ile His Lys Glu Tyr Leu Leu Ala Gly Ala Asp Ile Ile Glu Thr Asn	85		90	95	
act ttt agc agc act agt att gcc caa gct gac tat ggc ctt gaa cac					336
Thr Phe Ser Ser Thr Ser Ile Ala Gln Ala Asp Tyr Gly Leu Glu His	100		105	110	
ttg gcc tac cgg atg aac atg tgc tct gca gga gtg gcc aga aaa gct					384
Leu Ala Tyr Arg Met Asn Met Cys Ser Ala Gly Val Ala Arg Lys Ala	115		120	125	
gcc gag gag gta act ctc cag aca gga att aag agg ttt gtg gca ggg					432
Ala Glu Glu Val Thr Leu Gln Thr Gly Ile Lys Arg Phe Val Ala Gly	130		135	140	
gct ctg ggt ccg act aat aag aca ctc tct gtg tcc cca tct gtg gaa					480
Ala Leu Gly Pro Thr Asn Lys Thr Leu Ser Val Ser Pro Ser Val Glu	145		150	155	160
agg ccg gat tat agg aac atc aca ttt gat gag ctt gtt gaa gca tac					528
Arg Pro Asp Tyr Arg Asn Ile Thr Phe Asp Glu Leu Val Glu Ala Tyr	165		170	175	
caa gag cag gcc aaa gga ctt ctg gat ggc ggg gtt gat atc tta ctc					576
Gln Glu Gln Ala Lys Gly Leu Leu Asp Gly Gly Val Asp Ile Leu Leu	180		185	190	
att gaa act att ttt gat act gcc aat gcc aag gca gcc ttg ttt gca					624
Ile Glu Thr Ile Phe Asp Thr Ala Asn Ala Lys Ala Ala Leu Phe Ala	195		200	205	
ctc caa aat ctt ttt gag gag aaa tat gct ccc cgg cct atc ttt att					672
Leu Gln Asn Leu Phe Glu Glu Lys Tyr Ala Pro Arg Pro Ile Phe Ile	210		215	220	
tca ggg acg atc gtt gat aaa agt ggg cgg act ctt tcc gga cag aca					720
Ser Gly Thr Ile Val Asp Lys Ser Gly Arg Thr Leu Ser Gly Gln Thr	225		230	235	240
gga gag gga ttt gtc atc agc gtg tct cat gga gaa cca ctc tgc att					768
Gly Glu Gly Phe Val Ile Ser Val Ser His Gly Glu Pro Leu Cys Ile	245		250	255	
gga tta aat tgt gct ttg ggt gca gct gaa atg aga cct ttt att gaa					816
Gly Leu Asn Cys Ala Leu Gly Ala Ala Glu Met Arg Pro Phe Ile Glu	260		265	270	
ata att gga aaa tgt aca aca gcc tat gtc ctc tgt tat ccc aat gca					864
Ile Ile Gly Lys Cys Thr Thr Ala Tyr Val Leu Cys Tyr Pro Asn Ala	275		280	285	
ggg ctt ccc aac acc ttt ggt gac tat gat gaa acg cct tct atg atg					912
Gly Leu Pro Asn Thr Phe Gly Asp Tyr Asp Glu Thr Pro Ser Met Met	290		295	300	
gcc aag cac cta aag gat ttt gct atg gat ggc ttg gtc aat ata gtt					960
Ala Lys His Leu Lys Asp Phe Ala Met Asp Gly Leu Val Asn Ile Val	305		310	315	320

191

gga gga tgc tgt ggg tca aca cca gat cat atc agg gaa att gct gaa	1008
Gly Gly Cys Cys Gly Ser Thr Pro Asp His Ile Arg Glu Ile Ala Glu	
325 330 335	
gct gtg aaa aat tgt aag cct aga gtt cca cct gcc act gct ttt gaa	1056
Ala Val Lys Asn Cys Lys Pro Arg Val Pro Pro Ala Thr Ala Phe Glu	
340 345 350	
gga cat atg tta ctg tct ggt cta gag ccc ttc agg att gga ccg tac	1104
Gly His Met Leu Leu Ser Gly Leu Glu Pro Phe Arg Ile Gly Pro Tyr	
355 360 365	
acc aac ttt gtt aac att gga gag cgc tgt aat gtt gca gga tca agg	1152
Thr Asn Phe Val Asn Ile Gly Glu Arg Cys Asn Val Ala Gly Ser Arg	
370 375 380	
aag ttt gct aaa ctc atc atg gca gga aac tat gaa gaa gcc ttg tgt	1200
Lys Phe Ala Lys Leu Ile Met Ala Gly Asn Tyr Glu Glu Ala Leu Cys	
385 390 395 400	
gtt gcc aaa gtg cag gtg gaa atg gga gcc cag gtg ttg gat gtc aac	1248
Val Ala Lys Val Gln Val Glu Met Gly Ala Gln Val Leu Asp Val Asn	
405 410 415	
atg gat gat ggc atg cta gat ggt cca agt gca atg acc aga ttt tgc	1296
Met Asp Asp Gly Met Leu Asp Gly Pro Ser Ala Met Thr Arg Phe Cys	
420 425 430	
aac tta att gct tcc gag cca gac atc gca aag gta cct ttg tgc atc	1344
Asn Leu Ile Ala Ser Glu Pro Asp Ile Ala Lys Val Pro Leu Cys Ile	
435 440 445	
gac tcc tcc aat ttt gct gtg att gaa gct ggg tta aag tgc tgc caa	1392
Asp Ser Ser Asn Phe Ala Val Ile Glu Ala Gly Leu Lys Cys Cys Gln	
450 455 460	
ggg aag tgc att gtc aat agc att agt ctg aag gaa gga gag gac gac	1440
Gly Lys Cys Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Asp Asp	
465 470 475 480	
ttc ttg gag aag gcc agg aag att aaa aag tat gga gct gct atg gtg	1488
Phe Leu Glu Lys Ala Arg Lys Ile Lys Lys Tyr Gly Ala Ala Met Val	
485 490 495	
gtc atg gct ttt gat gaa gaa gga cag gca aca gaa aca gac aca aaa	1536
Val Met Ala Phe Asp Glu Glu Gly Gln Ala Thr Glu Thr Asp Thr Lys	
500 505 510	
atc aga gtg tgc acc cgg gcc tac cat ctg ctt gtg aaa aaa ctg ggc	1584
Ile Arg Val Cys Thr Arg Ala Tyr His Leu Leu Val Lys Lys Leu Gly	
515 520 525	
ttt aat cca aat gac att att ttt gac cct aat atc cta acc att ggg	1632
Phe Asn Pro Asn Asp Ile Ile Phe Asp Pro Asn Ile Leu Thr Ile Gly	
530 535 540	
act gga atg gag gaa cac aac ttg tat gcc att aat ttt atc cat gca	1680
Thr Gly Met Glu Glu His Asn Leu Tyr Ala Ile Asn Phe Ile His Ala	
545 550 555 560	
aca aaa gtc att aaa gaa aca tta cct gga gcc aga ata agt gga ggt	1728
Thr Lys Val Ile Lys Glu Thr Leu Pro Gly Ala Arg Ile Ser Gly Gly	

192															
565					570					575					
ctt tcc aac ttg tcc ttc tcc ttc cga gga atg gaa gcc att cga gaa	1776														
Leu Ser Asn Leu Ser Phe Ser Phe Arg Gly Met Glu Ala Ile Arg Glu															
580	585					590									
gca atg cat ggg gtt ttc ctt tac cat gca atc aag tct ggc atg gac	1824														
Ala Met His Gly Val Phe Leu Tyr His Ala Ile Lys Ser Gly Met Asp															
595	600					605									
atg ggg ata gtg aat gct gga aac ctc cct gtg tat gat gat atc cat	1872														
Met Gly Ile Val Asn Ala Gly Asn Leu Pro Val Tyr Asp Asp Ile His															
610	615					620									
aag gaa ctt ctg cag ctc tgt gaa gat ctc atc tgg aat aaa gac cct	1920														
Lys Glu Leu Leu Gln Leu Cys Glu Asp Leu Ile Trp Asn Lys Asp Pro															
625	630					635					640				
gag gcc act gag aag ctc tta cgt tat gcc cag act caa ggc aca gga	1968														
Glu Ala Thr Glu Lys Leu Leu Arg Tyr Ala Gln Thr Gln Gly Thr Gly															
645	650					655									
ggg aag aaa gtc att cag act gat gag tgg aga aat ggc cct gtc gaa	2016														
Gly Lys Lys Val Ile Gln Thr Asp Glu Trp Arg Asn Gly Pro Val Glu															
660	665					670									
gaa cgc ctt gag tat gcc ctt gtg aag ggc att gaa aaa cat att att	2064														
Glu Arg Leu Glu Tyr Ala Leu Val Lys Gly Ile Glu Lys His Ile Ile															
675	680					685									
gag gat act gag gaa gcc agg tta aac caa aaa aaa tat ccc cga cct	2112														
Glu Asp Thr Glu Glu Ala Arg Leu Asn Gln Lys Lys Tyr Pro Arg Pro															
690	695					700									
ctc aat ata att gaa gga ccc ctg atg aat gga atg aaa att gtt ggt	2160														
Leu Asn Ile Ile Glu Gly Pro Leu Met Asn Gly Met Lys Ile Val Gly															
705	710					715					720				
gat ctt ttt gga gct gga aaa atg ttt cta cct cag gtt ata aag tca	2208														
Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln Val Ile Lys Ser															
725	730					735									
gcc cgg gtt atg aag aag gct gtt ggc cac ctt atc cct ttc atg gaa	2256														
Ala Arg Val Met Lys Lys Ala Val Gly His Leu Ile Pro Phe Met Glu															
740	745					750									
aaa gaa aga gaa gaa acc aga gtg ctt aac ggc aca gta gaa gaa gag	2304														
Lys Glu Arg Glu Glu Thr Arg Val Leu Asn Gly Thr Val Glu Glu Glu															
755	760					765									
gac cct tac cag ggc acc atc gtg ctg gcc act gtt aaa ggc gac gtg	2352														
Asp Pro Tyr Gln Gly Thr Ile Val Leu Ala Thr Val Lys Gly Asp Val															
770	775					780									
cac gac ata ggc aag aac ata gtt gga gta gtc ctt ggc tgc aat aat	2400														
His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gly Cys Asn Asn															
785	790					795					800				
ttc cga gtt att gat tta gga gtc atg act cca tgt gat aag ata ctg	2448														
Phe Arg Val Ile Asp Leu Gly Val Met Thr Pro Cys Asp Lys Ile Leu															
805	810					815									

193

aaa gct gct ctt gac cac aaa gca gat ata att ggc ctg tca gga ctc	2496
Lys Ala Ala Leu Asp His Lys Ala Asp Ile Ile Gly Leu Ser Gly Leu	
820 825 830	
atc act cct tcc ctg gat gaa atg att ttt gtt gcc aag gaa atg gag	2544
Ile Thr Pro Ser Leu Asp Glu Met Ile Phe Val Ala Lys Glu Met Glu	
835 840 845	
aga tta gct ata agg att cca ttg ttg att gga gga gca acc act tca	2592
Arg Leu Ala Ile Arg Ile Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser	
850 855 860	
aaa acc cac aca gca gtt aaa ata gct ccg aga tac agt gca cct gta	2640
Lys Thr His Thr Ala Val Lys Ile Ala Pro Arg Tyr Ser Ala Pro Val	
865 870 875 880	
atc cat gtc ctg gac gcg tcc aag agt gtg gtg gtg tgt tcc cag ctg	2688
Ile His Val Leu Asp Ala Ser Lys Ser Val Val Val Cys Ser Gln Leu	
885 890 895	
tta gat gaa aat cta aag gat gaa tac ttt gag gaa atc atg gaa gaa	2736
Leu Asp Glu Asn Leu Lys Asp Glu Tyr Phe Glu Glu Ile Met Glu Glu	
900 905 910	
tat gaa gat att aga cag gac cat tat gag tct ctc aag gag agg aga	2784
Tyr Glu Asp Ile Arg Gln Asp His Tyr Glu Ser Leu Lys Glu Arg Arg	
915 920 925	
tac tta ccc tta agt caa gcc aga aaa agt ggt ttc caa atg gat tgg	2832
Tyr Leu Pro Leu Ser Gln Ala Arg Lys Ser Gly Phe Gln Met Asp Trp	
930 935 940	
ctg tct gaa cct cac cca gtg aag ccc acg ttt att ggg acc cag gtc	2880
Leu Ser Glu Pro His Pro Val Lys Pro Thr Phe Ile Gly Thr Gln Val	
945 950 955 960	
ttt gaa gac tat gac ctg cag aag ctg gtg gac tac att gac tgg aag	2928
Phe Glu Asp Tyr Asp Leu Gln Lys Leu Val Asp Tyr Ile Asp Trp Lys	
965 970 975	
cct ttc ttt gat gtc tgg cag ctc cgg ggc aag tac ccg aat cga ggc	2976
Pro Phe Phe Asp Val Trp Gln Leu Arg Gly Lys Tyr Pro Asn Arg Gly	
980 985 990	
ttt ccc aag ata ttt aac gac aaa aca gta ggt gga gag gcc agg aag	3024
Phe Pro Lys Ile Phe Asn Asp Lys Thr Val Gly Gly Glu Ala Arg Lys	
995 1000 1005	
gtc tac gat gat gcc cac aat atg ctg aac aca ctg att agt caa aag	3072
Val Tyr Asp Asp Ala His Asn Met Leu Asn Thr Leu Ile Ser Gln Lys	
1010 1015 1020	
aaa ctc cgg gcc cgg ggt gtg gtt ggg ttc tgg cca gca cag agt atc	3120
Lys Leu Arg Ala Arg Gly Val Val Gly Phe Trp Pro Ala Gln Ser Ile	
1025 1030 1035 1040	
caa gac gac att cac ctg tac gcg gag gct gct gtg ccc cag gct gca	3168
Gln Asp Asp Ile His Leu Tyr Ala Glu Ala Ala Val Pro Gln Ala Ala	
1045 1050 1055	
gag ccc ata gcc acc ttc tat ggg tta agg caa cag gct gag aag gac	3216
Glu Pro Ile Ala Thr Phe Tyr Gly Leu Arg Gln Gln Ala Glu Lys Asp	

194

1060	1065	1070	
tct gcc agc acg gag cca tac tac tgc ctc tca gac ttc atc gct ccc			3264
Ser Ala Ser Thr Glu Pro Tyr Tyr Cys Leu Ser Asp Phe Ile Ala Pro			
1075	1080	1085	
ttg cat tct ggc atc cgt gac tac ctg ggc ctg ttt gcc gtt gcc tgc			3312
Leu His Ser Gly Ile Arg Asp Tyr Leu Gly Leu Phe Ala Val Ala Cys			
1090	1095	1100	
ttt ggg gta gaa gag ctg agc aag gcc tat gag gat gat ggt gac gac			3360
Phe Gly Val Glu Glu Leu Ser Lys Ala Tyr Glu Asp Asp Gly Asp Asp			
1105	1110	1115	1120
tac agc agc atc atg gtc aag gcg ctg ggg gac cgg ctg gca gag gcc			3408
Tyr Ser Ser Ile Met Val Lys Ala Leu Gly Asp Arg Leu Ala Glu Ala			
1125	1130	1135	
ttt gca gaa gag ctc cat gaa aga gtt cgc cga gaa ctg tgg gcc tac			3456
Phe Ala Glu Glu Leu His Glu Arg Val Arg Arg Glu Leu Trp Ala Tyr			
1140	1145	1150	
tgt ggc agt gag cag ctg gac gtc gca gac ctg cgc agg ctg cgg tac			3504
Cys Gly Ser Glu Gln Leu Asp Val Ala Asp Leu Arg Arg Leu Arg Tyr			
1155	1160	1165	
aag ggc atc cgc ccg gct cct ggc tac ccc agc cag ccc gac cac acc			3552
Lys Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ser Gln Pro Asp His Thr			
1170	1175	1180	
gag aag ctc acc atg tgg aga ctt gca gac atc gag cag tct aca ggc			3600
Glu Lys Leu Thr Met Trp Arg Leu Ala Asp Ile Glu Gln Ser Thr Gly			
1185	1190	1195	1200
att agg tta aca gaa tca tta gca atg gca cct gct tca gca gtc tca			3648
Ile Arg Leu Thr Glu Ser Leu Ala Met Ala Pro Ala Ser Ala Val Ser			
1205	1210	1215	
ggc ctc tac ttc tcc aat ttg aag tcc aaa tat ttt gct gtg ggg aag			3696
Gly Leu Tyr Phe Ser Asn Leu Lys Ser Lys Tyr Phe Ala Val Gly Lys			
1220	1225	1230	
att tcc aag gat cag gtt gag gat tat gca ttg agg aag aac ata tct			3744
Ile Ser Lys Asp Gln Val Glu Asp Tyr Ala Leu Arg Lys Asn Ile Ser			
1235	1240	1245	
gtg gct gag gtt gag aaa tgg ctt gga ccc att ttg gga tat gat aca			3792
Val Ala Glu Val Glu Lys Trp Leu Gly Pro Ile Leu Gly Tyr Asp Thr			
1250	1255	1260	
gac taa			3798
Asp			
1265			

&lt;210&gt; 44

&lt;211&gt; 1265

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 44

Met Ser Pro Ala Leu Gln Asp Leu Ser Gln Pro Glu Gly Leu Lys Lys



195

1	5	10	15												
Thr	Leu	Arg	Asp	Glu	Ile	Asn	Ala	Ile	Leu	Gln	Lys	Arg	Ile	Met	Val
			20					25						30	
Leu	Asp	Gly	Gly	Met	Gly	Thr	Met	Ile	Gln	Arg	Glu	Lys	Leu	Asn	Glu
		35					40					45			
Glu	His	Phe	Arg	Gly	Gln	Glu	Phe	Lys	Asp	His	Ala	Arg	Pro	Leu	Lys
	50					55					60				
Gly	Asn	Asn	Asp	Ile	Leu	Ser	Ile	Thr	Gln	Pro	Asp	Val	Ile	Tyr	Gln
65					70					75					80
Ile	His	Lys	Glu	Tyr	Leu	Leu	Ala	Gly	Ala	Asp	Ile	Ile	Glu	Thr	Asn
				85					90						95
Thr	Phe	Ser	Ser	Thr	Ser	Ile	Ala	Gln	Ala	Asp	Tyr	Gly	Leu	Glu	His
			100					105						110	
Leu	Ala	Tyr	Arg	Met	Asn	Met	Cys	Ser	Ala	Gly	Val	Ala	Arg	Lys	Ala
		115					120					125			
Ala	Glu	Glu	Val	Thr	Leu	Gln	Thr	Gly	Ile	Lys	Arg	Phe	Val	Ala	Gly
	130					135					140				
Ala	Leu	Gly	Pro	Thr	Asn	Lys	Thr	Leu	Ser	Val	Ser	Pro	Ser	Val	Glu
145					150					155					160
Arg	Pro	Asp	Tyr	Arg	Asn	Ile	Thr	Phe	Asp	Glu	Leu	Val	Glu	Ala	Tyr
			165						170					175	
Gln	Glu	Gln	Ala	Lys	Gly	Leu	Leu	Asp	Gly	Gly	Val	Asp	Ile	Leu	Leu
			180					185					190		
Ile	Glu	Thr	Ile	Phe	Asp	Thr	Ala	Asn	Ala	Lys	Ala	Ala	Leu	Phe	Ala
	195						200						205		
Leu	Gln	Asn	Leu	Phe	Glu	Glu	Lys	Tyr	Ala	Pro	Arg	Pro	Ile	Phe	Ile
	210						215					220			
Ser	Gly	Thr	Ile	Val	Asp	Lys	Ser	Gly	Arg	Thr	Leu	Ser	Gly	Gln	Thr
225					230					235					240
Gly	Glu	Gly	Phe	Val	Ile	Ser	Val	Ser	His	Gly	Glu	Pro	Leu	Cys	Ile
			245						250					255	
Gly	Leu	Asn	Cys	Ala	Leu	Gly	Ala	Ala	Glu	Met	Arg	Pro	Phe	Ile	Glu
		260						265					270		
Ile	Ile	Gly	Lys	Cys	Thr	Thr	Ala	Tyr	Val	Leu	Cys	Tyr	Pro	Asn	Ala
		275					280						285		
Gly	Leu	Pro	Asn	Thr	Phe	Gly	Asp	Tyr	Asp	Glu	Thr	Pro	Ser	Met	Met
	290						295					300			
Ala	Lys	His	Leu	Lys	Asp	Phe	Ala	Met	Asp	Gly	Leu	Val	Asn	Ile	Val
305					310					315					320
Gly	Gly	Cys	Cys	Gly	Ser	Thr	Pro	Asp	His	Ile	Arg	Glu	Ile	Ala	Glu
			325						330						335

196

Ala Val Lys Asn Cys Lys Pro Arg Val Pro Pro Ala Thr Ala Phe Glu  
 340 345 350  
 Gly His Met Leu Leu Ser Gly Leu Glu Pro Phe Arg Ile Gly Pro Tyr  
 355 360 365  
 Thr Asn Phe Val Asn Ile Gly Glu Arg Cys Asn Val Ala Gly Ser Arg  
 370 375 380  
 Lys Phe Ala Lys Leu Ile Met Ala Gly Asn Tyr Glu Glu Ala Leu Cys  
 385 390 395 400  
 Val Ala Lys Val Gln Val Glu Met Gly Ala Gln Val Leu Asp Val Asn  
 405 410 415  
 Met Asp Asp Gly Met Leu Asp Gly Pro Ser Ala Met Thr Arg Phe Cys  
 420 425 430  
 Asn Leu Ile Ala Ser Glu Pro Asp Ile Ala Lys Val Pro Leu Cys Ile  
 435 440 445  
 Asp Ser Ser Asn Phe Ala Val Ile Glu Ala Gly Leu Lys Cys Cys Gln  
 450 455 460  
 Gly Lys Cys Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Asp Asp  
 465 470 475 480  
 Phe Leu Glu Lys Ala Arg Lys Ile Lys Lys Tyr Gly Ala Ala Met Val  
 485 490 495  
 Val Met Ala Phe Asp Glu Glu Gly Gln Ala Thr Glu Thr Asp Thr Lys  
 500 505 510  
 Ile Arg Val Cys Thr Arg Ala Tyr His Leu Leu Val Lys Lys Leu Gly  
 515 520 525  
 Phe Asn Pro Asn Asp Ile Ile Phe Asp Pro Asn Ile Leu Thr Ile Gly  
 530 535 540  
 Thr Gly Met Glu Glu His Asn Leu Tyr Ala Ile Asn Phe Ile His Ala  
 545 550 555 560  
 Thr Lys Val Ile Lys Glu Thr Leu Pro Gly Ala Arg Ile Ser Gly Gly  
 565 570 575  
 Leu Ser Asn Leu Ser Phe Ser Phe Arg Gly Met Glu Ala Ile Arg Glu  
 580 585 590  
 Ala Met His Gly Val Phe Leu Tyr His Ala Ile Lys Ser Gly Met Asp  
 595 600 605  
 Met Gly Ile Val Asn Ala Gly Asn Leu Pro Val Tyr Asp Asp Ile His  
 610 615 620  
 Lys Glu Leu Leu Gln Leu Cys Glu Asp Leu Ile Trp Asn Lys Asp Pro  
 625 630 635 640  
 Glu Ala Thr Glu Lys Leu Leu Arg Tyr Ala Gln Thr Gln Gly Thr Gly  
 645 650 655  
 Gly Lys Lys Val Ile Gln Thr Asp Glu Trp Arg Asn Gly Pro Val Glu  
 660 665 670

197

Glu Arg Leu Glu Tyr Ala Leu Val Lys Gly Ile Glu Lys His Ile Ile  
 675 680 685  
 Glu Asp Thr Glu Glu Ala Arg Leu Asn Gln Lys Lys Tyr Pro Arg Pro  
 690 695 700  
 Leu Asn Ile Ile Glu Gly Pro Leu Met Asn Gly Met Lys Ile Val Gly  
 705 710 715 720  
 Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln Val Ile Lys Ser  
 725 730 735  
 Ala Arg Val Met Lys Lys Ala Val Gly His Leu Ile Pro Phe Met Glu  
 740 745 750  
 Lys Glu Arg Glu Glu Thr Arg Val Leu Asn Gly Thr Val Glu Glu Glu  
 755 760 765  
 Asp Pro Tyr Gln Gly Thr Ile Val Leu Ala Thr Val Lys Gly Asp Val  
 770 775 780  
 His Asp Ile Gly Lys Asn Ile Val Gly Val Val Leu Gly Cys Asn Asn  
 785 790 795 800  
 Phe Arg Val Ile Asp Leu Gly Val Met Thr Pro Cys Asp Lys Ile Leu  
 805 810 815  
 Lys Ala Ala Leu Asp His Lys Ala Asp Ile Ile Gly Leu Ser Gly Leu  
 820 825 830  
 Ile Thr Pro Ser Leu Asp Glu Met Ile Phe Val Ala Lys Glu Met Glu  
 835 840 845  
 Arg Leu Ala Ile Arg Ile Pro Leu Leu Ile Gly Gly Ala Thr Thr Ser  
 850 855 860  
 Lys Thr His Thr Ala Val Lys Ile Ala Pro Arg Tyr Ser Ala Pro Val  
 865 870 875 880  
 Ile His Val Leu Asp Ala Ser Lys Ser Val Val Val Cys Ser Gln Leu  
 885 890 895  
 Leu Asp Glu Asn Leu Lys Asp Glu Tyr Phe Glu Glu Ile Met Glu Glu  
 900 905 910  
 Tyr Glu Asp Ile Arg Gln Asp His Tyr Glu Ser Leu Lys Glu Arg Arg  
 915 920 925  
 Tyr Leu Pro Leu Ser Gln Ala Arg Lys Ser Gly Phe Gln Met Asp Trp  
 930 935 940  
 Leu Ser Glu Pro His Pro Val Lys Pro Thr Phe Ile Gly Thr Gln Val  
 945 950 955 960  
 Phe Glu Asp Tyr Asp Leu Gln Lys Leu Val Asp Tyr Ile Asp Trp Lys  
 965 970 975  
 Pro Phe Phe Asp Val Trp Gln Leu Arg Gly Lys Tyr Pro Asn Arg Gly  
 980 985 990  
 Phe Pro Lys Ile Phe Asn Asp Lys Thr Val Gly Gly Glu Ala Arg Lys

198

995	1000	1005	
Val Tyr Asp Asp Ala His Asn Met Leu Asn Thr Leu Ile Ser Gln Lys			
1010	1015	1020	
Lys Leu Arg Ala Arg Gly Val Val Gly Phe Trp Pro Ala Gln Ser Ile			
1025	1030	1035	1040
Gln Asp Asp Ile His Leu Tyr Ala Glu Ala Ala Val Pro Gln Ala Ala			
	1045	1050	1055
Glu Pro Ile Ala Thr Phe Tyr Gly Leu Arg Gln Gln Ala Glu Lys Asp			
	1060	1065	1070
Ser Ala Ser Thr Glu Pro Tyr Tyr Cys Leu Ser Asp Phe Ile Ala Pro			
	1075	1080	1085
Leu His Ser Gly Ile Arg Asp Tyr Leu Gly Leu Phe Ala Val Ala Cys			
	1090	1095	1100
Phe Gly Val Glu Glu Leu Ser Lys Ala Tyr Glu Asp Asp Gly Asp Asp			
1105	1110	1115	1120
Tyr Ser Ser Ile Met Val Lys Ala Leu Gly Asp Arg Leu Ala Glu Ala			
	1125	1130	1135
Phe Ala Glu Glu Leu His Glu Arg Val Arg Arg Glu Leu Trp Ala Tyr			
	1140	1145	1150
Cys Gly Ser Glu Gln Leu Asp Val Ala Asp Leu Arg Arg Leu Arg Tyr			
	1155	1160	1165
Lys Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ser Gln Pro Asp His Thr			
	1170	1175	1180
Glu Lys Leu Thr Met Trp Arg Leu Ala Asp Ile Glu Gln Ser Thr Gly			
1185	1190	1195	1200
Ile Arg Leu Thr Glu Ser Leu Ala Met Ala Pro Ala Ser Ala Val Ser			
	1205	1210	1215
Gly Leu Tyr Phe Ser Asn Leu Lys Ser Lys Tyr Phe Ala Val Gly Lys			
	1220	1225	1230
Ile Ser Lys Asp Gln Val Glu Asp Tyr Ala Leu Arg Lys Asn Ile Ser			
	1235	1240	1245
Val Ala Glu Val Glu Lys Trp Leu Gly Pro Ile Leu Gly Tyr Asp Thr			
	1250	1255	1260
Asp			
1265			

<210> 45  
 <211> 3681  
 <212> DNA  
 <213> Vibrio fisheri

<220>  
 <221> CDS  
 <222> (1) .. (3678)

&lt;223&gt; AB039955

&lt;400&gt; 45

gtg gca gga agc aat ata aaa gta caa ata gaa aag caa ctt tca gag	48
Val Ala Gly Ser Asn Ile Lys Val Gln Ile Glu Lys Gln Leu Ser Glu	
1 5 10 15	
cga att tta ttg att gat ggt ggt atg ggc acc atg att caa ggt tat	96
Arg Ile Leu Leu Ile Asp Gly Gly Met Gly Thr Met Ile Gln Gly Tyr	
20 25 30	
aag ttt gaa gag aaa gat tat aga ggg gga cgc ttt aat caa tgg cat	144
Lys Phe Glu Glu Lys Asp Tyr Arg Gly Gly Arg Phe Asn Gln Trp His	
35 40 45	
tgt gat ctt aaa ggt aac aat gat tta tta gtt ctt tca caa cca caa	192
Cys Asp Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Ser Gln Pro Gln	
50 55 60	
att ata aga gat ata cac gaa gcc tat tta gaa gct ggt gct gat atc	240
Ile Ile Arg Asp Ile His Glu Ala Tyr Leu Glu Ala Gly Ala Asp Ile	
65 70 75 80	
ctt gaa act aat acc ttt aat gca aca act att gct atg gct gat tat	288
Leu Glu Thr Asn Thr Phe Asn Ala Thr Thr Ile Ala Met Ala Asp Tyr	
85 90 95	
gat atg gaa agc ctt agt gaa gag att aac ttt gaa gca gca aag ctt	336
Asp Met Glu Ser Leu Ser Glu Glu Ile Asn Phe Glu Ala Ala Lys Leu	
100 105 110	
gct cgt gaa gtt gca gat aaa tgg aca gaa aaa aca cca aac aaa cct	384
Ala Arg Glu Val Ala Asp Lys Trp Thr Glu Lys Thr Pro Asn Lys Pro	
115 120 125	
cgc tat gta gca gga gtg ctt gga cca aca aat cga act tgt tct att	432
Arg Tyr Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Cys Ser Ile	
130 135 140	
tct cca gac gta aat gac cct ggc ttt cgt aat gta tcg ttt gat gaa	480
Ser Pro Asp Val Asn Asp Pro Gly Phe Arg Asn Val Ser Phe Asp Glu	
145 150 155 160	
tta gtc gaa gct tat tca gag tca act cga gca ctt att aga ggt ggt	528
Leu Val Glu Ala Tyr Ser Glu Ser Thr Arg Ala Leu Ile Arg Gly Gly	
165 170 175	
tca gat ctt atc ctc atc gaa act ata ttt gat aca tta aat gct aaa	576
Ser Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys	
180 185 190	
gcg tgt tct ttt gct gtt gaa tct gtt ttt gaa gag ctt ggt att act	624
Ala Cys Ser Phe Ala Val Glu Ser Val Phe Glu Glu Leu Gly Ile Thr	
195 200 205	
ttg cct gtt atg att tca ggg acc att acc gat gca tca gga aga aca	672
Leu Pro Val Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr	
210 215 220	
tta tcg ggg caa aca aca gaa gct ttt tat aat gca tta aga cat gta	720
Leu Ser Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ala Leu Arg His Val	
225 230 235 240	

## 200

aaa cct att tct ttt ggt ctt aac tgt gca ctt ggt cct gat gaa tta	768
Lys Pro Ile Ser Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu	
245 250 255	
cgt gaa tat gta agc gag ctt tca cgt att tct gaa tgt tat gtt tct	816
Arg Glu Tyr Val Ser Glu Leu Ser Arg Ile Ser Glu Cys Tyr Val Ser	
260 265 270	
gcg cac cca aac gct ggt ttg cct aat gca ttt ggt gag tat gat tta	864
Ala His Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu	
275 280 285	
tct ccc gaa gat atg gct gag cat gtt gcg gaa tgg gca agc agc gga	912
Ser Pro Glu Asp Met Ala Glu His Val Ala Glu Trp Ala Ser Ser Gly	
290 295 300	
ttt tta aat ctt att ggt ggg tgt tgt ggc acc act cct gaa cat att	960
Phe Leu Asn Leu Ile Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile	
305 310 315 320	
cgt caa atg gct tta gtt gtt gaa ggt gtg aaa cct cga caa tta cct	1008
Arg Gln Met Ala Leu Val Val Glu Gly Val Lys Pro Arg Gln Leu Pro	
325 330 335	
gaa tta ccc gta gct tgt cgt ctt tcc gga tta gag cct tta aca ata	1056
Glu Leu Pro Val Ala Cys Arg Leu Ser Gly Leu Glu Pro Leu Thr Ile	
340 345 350	
gaa aaa gat tct ttg ttt att aat gtt ggt gaa cgt aca aat gtt act	1104
Glu Lys Asp Ser Leu Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr	
355 360 365	
gga tct gca cgt ttt aaa cgc tta att aaa gaa gag ctt tat gac gaa	1152
Gly Ser Ala Arg Phe Lys Arg Leu Ile Lys Glu Glu Leu Tyr Asp Glu	
370 375 380	
gca cta agt gtt gct caa gag caa gtt gaa aac ggt gct caa att atc	1200
Ala Leu Ser Val Ala Gln Glu Gln Val Glu Asn Gly Ala Gln Ile Ile	
385 390 395 400	
gat atc aac atg gat gaa ggc atg ctt gat gct gaa gca tgt atg gtt	1248
Asp Ile Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Cys Met Val	
405 410 415	
cgt ttt tta aat ctt tgt gca tca gaa cct gaa ata tct aaa gta cca	1296
Arg Phe Leu Asn Leu Cys Ala Ser Glu Pro Glu Ile Ser Lys Val Pro	
420 425 430	
gtg atg gtt gat tct tct aaa tgg gaa gta att gaa gct gga tta aag	1344
Val Met Val Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys	
435 440 445	
tgt att caa ggt aag ggg ata gtt aat tca atc tct tta aag gaa ggc	1392
Cys Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Leu Lys Glu Gly	
450 455 460	
aaa gaa aag ttt gta cat caa gcc aag tta ata cgt cgt tat ggt gct	1440
Lys Glu Lys Phe Val His Gln Ala Lys Leu Ile Arg Arg Tyr Gly Ala	
465 470 475 480	
gca gtg atc gtt atg gct ttt gat gaa gtt ggc caa gcg gac act cgg	1488

201

Ala Val Ile Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Arg	
485 490 495	
gag cgt aaa att gaa att tgt acc aat gcc tac aat att tta gtt gat	1536
Glu Arg Lys Ile Glu Ile Cys Thr Asn Ala Tyr Asn Ile Leu Val Asp	
500 505 510	
gaa gtt ggc ttc cca cct gaa gat att att ttt gac cct aat att ttt	1584
Glu Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe	
515 520 525	
gcg gtt gct aca ggt atc gat gaa cat aat aac tat gca gta gac ttt	1632
Ala Val Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Val Asp Phe	
530 535 540	
att gaa gcc gtt ggt gat ata aag cga acg ctt cct cat gca atg att	1680
Ile Glu Ala Val Gly Asp Ile Lys Arg Thr Leu Pro His Ala Met Ile	
545 550 555 560	
tca ggt ggt gtt tct aac gtc tct ttt tct ttc cgt gga aat aac tac	1728
Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn Tyr	
565 570 575	
gtt cgt gaa gct atc cat gcc gta ttt tta tat cac tgt ttt aaa aat	1776
Val Arg Glu Ala Ile His Ala Val Phe Leu Tyr His Cys Phe Lys Asn	
580 585 590	
ggt atg gat atg ggc atc gta aat gcg ggg cag ctg gaa ata tat gat	1824
Gly Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr Asp	
595 600 605	
aac gta cca gaa gat ctg cgt gaa gcg gtt gaa gat gtg gta ttg aat	1872
Asn Val Pro Glu Asp Leu Arg Glu Ala Val Glu Asp Val Val Leu Asn	
610 615 620	
cgt cga gat gat tct acg gag cgt tta ctt gat att gca act gag tat	1920
Arg Arg Asp Asp Ser Thr Glu Arg Leu Leu Asp Ile Ala Thr Glu Tyr	
625 630 635 640	
tta gaa cga gct gtt ggt aaa gtt gaa gat aaa tct gct tta gag tgg	1968
Leu Glu Arg Ala Val Gly Lys Val Glu Asp Lys Ser Ala Leu Glu Trp	
645 650 655	
cgt gac tgg cct gtt gaa aaa cgt ctt gag cat tct cta gtg aag ggg	2016
Arg Asp Trp Pro Val Glu Lys Arg Leu Glu His Ser Leu Val Lys Gly	
660 665 670	
ata aca gag ttt att gtc gaa gat aca gaa gaa gca cga atc aat gca	2064
Ile Thr Glu Phe Ile Val Glu Asp Thr Glu Glu Ala Arg Ile Asn Ala	
675 680 685	
gaa aga cca ata gag gta att gaa ggg cca ttg atg gac gga atg aac	2112
Glu Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn	
690 695 700	
gtc gtt ggt gat ctt ttt ggg gaa gga aaa atg ttc ctt ccc caa gta	2160
Val Val Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val	
705 710 715 720	
gta aag tct gct cgt gta atg aaa caa gct gtt gct cat tta gaa ccg	2208
Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Glu Pro	
725 730 735	

## 202

ttt att aat gcg tct aaa gaa gtt gga gca aca aac ggt aaa ata ctt Phe Ile Asn Ala Ser Lys Glu Val Gly Ala Thr Asn Gly Lys Ile Leu 740 745 750	2256
tta gca aca gta aaa ggt gat gtt cat gat att ggt aag aat atc gtt Leu Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val 755 760 765	2304
ggc gtg gtt tta cag tgt aat aac tat gaa ata att gat ctt ggt gtc Gly Val Val Leu Gln Cys Asn Asn Tyr Glu Ile Ile Asp Leu Gly Val 770 775 780	2352
atg gtc tct tgt gaa act atc tta aaa gta gcc aaa gaa gaa aat gta Met Val Ser Cys Glu Thr Ile Leu Lys Val Ala Lys Glu Glu Asn Val 785 790 795 800	2400
gac atc att ggt tta tct gga tta ata aca cca tca tta gat gaa atg Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met 805 810 815	2448
gtc cat gtt gct aaa gag atg gaa cga caa ggg ttt gat tta cca ttg Val His Val Ala Lys Glu Met Glu Arg Gln Gly Phe Asp Leu Pro Leu 820 825 830	2496
ttg att ggt gga gca aca act tca aaa gca cat aca gcg gta aaa att Leu Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile 835 840 845	2544
gaa caa aac tat tct caa cct gtt gtg tac gtt aat aat gct tct cga Glu Gln Asn Tyr Ser Gln Pro Val Val Tyr Val Asn Asn Ala Ser Arg 850 855 860	2592
gct gta ggt gta tgt act tca tta ctt tca aat gaa cta aaa cct tct Ala Val Gly Val Cys Thr Ser Leu Leu Ser Asn Glu Leu Lys Pro Ser 865 870 875 880	2640
ttt gtt gag aag cta gat att gat tac gaa cgt gtt aga gag cag cat Phe Val Glu Lys Leu Asp Ile Asp Tyr Glu Arg Val Arg Glu Gln His 885 890 895	2688
agt cgt aaa caa ccg cga act aag cct gtg act tta gag gtt gct cga Ser Arg Lys Gln Pro Arg Thr Lys Pro Val Thr Leu Glu Val Ala Arg 900 905 910	2736
gcg aat aaa gtc gct att gac tgg gct tct tat aca cct cct gtc cca Ala Asn Lys Val Ala Ile Asp Trp Ala Ser Tyr Thr Pro Pro Val Pro 915 920 925	2784
cta aag cct ggt gta cat ata ttt gat aac ttt gat gtt tca aca ttg Leu Lys Pro Gly Val His Ile Phe Asp Asn Phe Asp Val Ser Thr Leu 930 935 940	2832
cgt aat tat att gat tgg acc cca ttt ttt atg acg tgg tct ctt gtt Arg Asn Tyr Ile Asp Trp Thr Pro Phe Phe Met Thr Trp Ser Leu Val 945 950 955 960	2880
gga aaa tac ccg aag atc tta gag cat gaa gaa gtt ggt gaa gaa gcc Gly Lys Tyr Pro Lys Ile Leu Glu His Glu Glu Val Gly Glu Glu Ala 965 970 975	2928
aaa cga tta ttt aaa gat gca aat gat cta tta gat cga gtt gaa aaa	2976



## 203

Lys Arg Leu Phe Lys Asp Ala Asn Asp Leu Leu Asp Arg Val Glu Lys  
 980 985 990

gaa ggg tta ctt aaa gcc cgt gga atg tgt gcg cta ttt cca gct tcc 3024  
 Glu Gly Leu Leu Lys Ala Arg Gly Met Cys Ala Leu Phe Pro Ala Ser  
 995 1000 1005

agt gtt ggt gat gat att gaa gta tat act gat gaa tca cgc act aca 3072  
 Ser Val Gly Asp Asp Ile Glu Val Tyr Thr Asp Glu Ser Arg Thr Thr  
 1010 1015 1020

gtt gca aaa gta ctt cat aat ttg cga caa caa acg gag aag ccg aaa 3120  
 Val Ala Lys Val Leu His Asn Leu Arg Gln Gln Thr Glu Lys Pro Lys  
 1025 1030 1035 1040

ggt ttt aat tat tgt tta tct gat tat ata gca ccc aaa gag tcg ggt 3168  
 Gly Phe Asn Tyr Cys Leu Ser Asp Tyr Ile Ala Pro Lys Glu Ser Gly  
 1045 1050 1055

aaa aat gat tgg atc ggt ggt ttt gct gta act ggt ggt att ggt gag 3216  
 Lys Asn Asp Trp Ile Gly Gly Phe Ala Val Thr Gly Gly Ile Gly Glu  
 1060 1065 1070

cgt gaa cta gct gat gaa tat aaa gca aat ggt gat gat tat aac gct 3264  
 Arg Glu Leu Ala Asp Glu Tyr Lys Ala Asn Gly Asp Asp Tyr Asn Ala  
 1075 1080 1085

atc atg att caa gcg gtg gct gat cgt cta gct gaa gct ttt gct gaa 3312  
 Ile Met Ile Gln Ala Val Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu  
 1090 1095 1100

tat tta cat gaa aaa gta cgt aag gaa att tgg ggt tac tct cct aat 3360  
 Tyr Leu His Glu Lys Val Arg Lys Glu Ile Trp Gly Tyr Ser Pro Asn  
 1105 1110 1115 1120

gag acg ctt tca aat gat gat tta atc cgt gaa aaa tac caa ggc att 3408  
 Glu Thr Leu Ser Asn Asp Asp Leu Ile Arg Glu Lys Tyr Gln Gly Ile  
 1125 1130 1135

cgt cct gct cct ggt tac cca gct tgt cct gaa cat aca gaa aaa ggg 3456  
 Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Gly  
 1140 1145 1150

gct tta tgg gag tta atg aat gtt gaa gaa tct att gga atg tct tta 3504  
 Ala Leu Trp Glu Leu Met Asn Val Glu Glu Ser Ile Gly Met Ser Leu  
 1155 1160 1165

aca tca agc tat gca atg tgg ccc ggt gca tct gtg tca gga atg tat 3552  
 Thr Ser Ser Tyr Ala Met Trp Pro Gly Ala Ser Val Ser Gly Met Tyr  
 1170 1175 1180

ttt tca cac cca gat tct cgt tat ttt gcg att gct cag att cag caa 3600  
 Phe Ser His Pro Asp Ser Arg Tyr Phe Ala Ile Ala Gln Ile Gln Gln  
 1185 1190 1195 1200

gat caa gcc gaa agc tat gcc gat cgt aaa ggt tgg aat atg ctt gaa 3648  
 Asp Gln Ala Glu Ser Tyr Ala Asp Arg Lys Gly Trp Asn Met Leu Glu  
 1205 1210 1215

gct gag aag tgg tta ggt cca aat ttg aat taa 3681  
 Ala Glu Lys Trp Leu Gly Pro Asn Leu Asn  
 1220 1225

<210> 46  
 <211> 1226  
 <212> PRT  
 <213> *Vibrio fisheri*

<400> 46

Val Ala Gly Ser Asn Ile Lys Val Gln Ile Glu Lys Gln Leu Ser Glu  
 1 5 10 15  
 Arg Ile Leu Leu Ile Asp Gly Gly Met Gly Thr Met Ile Gln Gly Tyr  
 20 25 30  
 Lys Phe Glu Glu Lys Asp Tyr Arg Gly Gly Arg Phe Asn Gln Trp His  
 35 40 45  
 Cys Asp Leu Lys Gly Asn Asn Asp Leu Leu Val Leu Ser Gln Pro Gln  
 50 55 60  
 Ile Ile Arg Asp Ile His Glu Ala Tyr Leu Glu Ala Gly Ala Asp Ile  
 65 70 75 80  
 Leu Glu Thr Asn Thr Phe Asn Ala Thr Thr Ile Ala Met Ala Asp Tyr  
 85 90 95  
 Asp Met Glu Ser Leu Ser Glu Glu Ile Asn Phe Glu Ala Ala Lys Leu  
 100 105 110  
 Ala Arg Glu Val Ala Asp Lys Trp Thr Glu Lys Thr Pro Asn Lys Pro  
 115 120 125  
 Arg Tyr Val Ala Gly Val Leu Gly Pro Thr Asn Arg Thr Cys Ser Ile  
 130 135 140  
 Ser Pro Asp Val Asn Asp Pro Gly Phe Arg Asn Val Ser Phe Asp Glu  
 145 150 155 160  
 Leu Val Glu Ala Tyr Ser Glu Ser Thr Arg Ala Leu Ile Arg Gly Gly  
 165 170 175  
 Ser Asp Leu Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys  
 180 185 190  
 Ala Cys Ser Phe Ala Val Glu Ser Val Phe Glu Glu Leu Gly Ile Thr  
 195 200 205  
 Leu Pro Val Met Ile Ser Gly Thr Ile Thr Asp Ala Ser Gly Arg Thr  
 210 215 220  
 Leu Ser Gly Gln Thr Thr Glu Ala Phe Tyr Asn Ala Leu Arg His Val  
 225 230 235 240  
 Lys Pro Ile Ser Phe Gly Leu Asn Cys Ala Leu Gly Pro Asp Glu Leu  
 245 250 255  
 Arg Glu Tyr Val Ser Glu Leu Ser Arg Ile Ser Glu Cys Tyr Val Ser  
 260 265 270  
 Ala His Pro Asn Ala Gly Leu Pro Asn Ala Phe Gly Glu Tyr Asp Leu  
 275 280 285

205

Ser Pro Glu Asp Met Ala Glu His Val Ala Glu Trp Ala Ser Ser Gly  
 290 295 300  
 Phe Leu Asn Leu Ile Gly Gly Cys Cys Gly Thr Thr Pro Glu His Ile  
 305 310 315 320  
 Arg Gln Met Ala Leu Val Val Glu Gly Val Lys Pro Arg Gln Leu Pro  
 325 330 335  
 Glu Leu Pro Val Ala Cys Arg Leu Ser Gly Leu Glu Pro Leu Thr Ile  
 340 345 350  
 Glu Lys Asp Ser Leu Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr  
 355 360 365  
 Gly Ser Ala Arg Phe Lys Arg Leu Ile Lys Glu Glu Leu Tyr Asp Glu  
 370 375 380  
 Ala Leu Ser Val Ala Gln Glu Gln Val Glu Asn Gly Ala Gln Ile Ile  
 385 390 395 400  
 Asp Ile Asn Met Asp Glu Gly Met Leu Asp Ala Glu Ala Cys Met Val  
 405 410 415  
 Arg Phe Leu Asn Leu Cys Ala Ser Glu Pro Glu Ile Ser Lys Val Pro  
 420 425 430  
 Val Met Val Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys  
 435 440 445  
 Cys Ile Gln Gly Lys Gly Ile Val Asn Ser Ile Ser Leu Lys Glu Gly  
 450 455 460  
 Lys Glu Lys Phe Val His Gln Ala Lys Leu Ile Arg Arg Tyr Gly Ala  
 465 470 475 480  
 Ala Val Ile Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Arg  
 485 490 495  
 Glu Arg Lys Ile Glu Ile Cys Thr Asn Ala Tyr Asn Ile Leu Val Asp  
 500 505 510  
 Glu Val Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe  
 515 520 525  
 Ala Val Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Val Asp Phe  
 530 535 540  
 Ile Glu Ala Val Gly Asp Ile Lys Arg Thr Leu Pro His Ala Met Ile  
 545 550 555 560  
 Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asn Tyr  
 565 570 575  
 Val Arg Glu Ala Ile His Ala Val Phe Leu Tyr His Cys Phe Lys Asn  
 580 585 590  
 Gly Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Glu Ile Tyr Asp  
 595 600 605  
 Asn Val Pro Glu Asp Leu Arg Glu Ala Val Glu Asp Val Val Leu Asn  
 610 615 620

206

Arg Arg Asp Asp Ser Thr Glu Arg Leu Leu Asp Ile Ala Thr Glu Tyr  
 625 630 635 640  
 Leu Glu Arg Ala Val Gly Lys Val Glu Asp Lys Ser Ala Leu Glu Trp  
 645 650 655  
 Arg Asp Trp Pro Val Glu Lys Arg Leu Glu His Ser Leu Val Lys Gly  
 660 665 670  
 Ile Thr Glu Phe Ile Val Glu Asp Thr Glu Glu Ala Arg Ile Asn Ala  
 675 680 685  
 Glu Arg Pro Ile Glu Val Ile Glu Gly Pro Leu Met Asp Gly Met Asn  
 690 695 700  
 Val Val Gly Asp Leu Phe Gly Glu Gly Lys Met Phe Leu Pro Gln Val  
 705 710 715 720  
 Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Glu Pro  
 725 730 735  
 Phe Ile Asn Ala Ser Lys Glu Val Gly Ala Thr Asn Gly Lys Ile Leu  
 740 745 750  
 Leu Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val  
 755 760 765  
 Gly Val Val Leu Gln Cys Asn Asn Tyr Glu Ile Ile Asp Leu Gly Val  
 770 775 780  
 Met Val Ser Cys Glu Thr Ile Leu Lys Val Ala Lys Glu Glu Asn Val  
 785 790 795 800  
 Asp Ile Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu Met  
 805 810 815  
 Val His Val Ala Lys Glu Met Glu Arg Gln Gly Phe Asp Leu Pro Leu  
 820 825 830  
 Leu Ile Gly Gly Ala Thr Thr Ser Lys Ala His Thr Ala Val Lys Ile  
 835 840 845  
 Glu Gln Asn Tyr Ser Gln Pro Val Val Tyr Val Asn Asn Ala Ser Arg  
 850 855 860  
 Ala Val Gly Val Cys Thr Ser Leu Leu Ser Asn Glu Leu Lys Pro Ser  
 865 870 875 880  
 Phe Val Glu Lys Leu Asp Ile Asp Tyr Glu Arg Val Arg Glu Gln His  
 885 890 895  
 Ser Arg Lys Gln Pro Arg Thr Lys Pro Val Thr Leu Glu Val Ala Arg  
 900 905 910  
 Ala Asn Lys Val Ala Ile Asp Trp Ala Ser Tyr Thr Pro Pro Val Pro  
 915 920 925  
 Leu Lys Pro Gly Val His Ile Phe Asp Asn Phe Asp Val Ser Thr Leu  
 930 935 940  
 Arg Asn Tyr Ile Asp Trp Thr Pro Phe Phe Met Thr Trp Ser Leu Val

207

945	950	955	960
Gly Lys Tyr Pro Lys Ile Leu Glu His Glu Glu Val Gly Glu Glu Ala	965	970	975
Lys Arg Leu Phe Lys Asp Ala Asn Asp Leu Leu Asp Arg Val Glu Lys	980	985	990
Glu Gly Leu Leu Lys Ala Arg Gly Met Cys Ala Leu Phe Pro Ala Ser	995	1000	1005
Ser Val Gly Asp Asp Ile Glu Val Tyr Thr Asp Glu Ser Arg Thr Thr	1010	1015	1020
Val Ala Lys Val Leu His Asn Leu Arg Gln Gln Thr Glu Lys Pro Lys	1025	1030	1035
Gly Phe Asn Tyr Cys Leu Ser Asp Tyr Ile Ala Pro Lys Glu Ser Gly	1045	1050	1055
Lys Asn Asp Trp Ile Gly Gly Phe Ala Val Thr Gly Gly Ile Gly Glu	1060	1065	1070
Arg Glu Leu Ala Asp Glu Tyr Lys Ala Asn Gly Asp Asp Tyr Asn Ala	1075	1080	1085
Ile Met Ile Gln Ala Val Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu	1090	1095	1100
Tyr Leu His Glu Lys Val Arg Lys Glu Ile Trp Gly Tyr Ser Pro Asn	1105	1110	1115
Glu Thr Leu Ser Asn Asp Asp Leu Ile Arg Glu Lys Tyr Gln Gly Ile	1125	1130	1135
Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Glu Lys Gly	1140	1145	1150
Ala Leu Trp Glu Leu Met Asn Val Glu Glu Ser Ile Gly Met Ser Leu	1155	1160	1165
Thr Ser Ser Tyr Ala Met Trp Pro Gly Ala Ser Val Ser Gly Met Tyr	1170	1175	1180
Phe Ser His Pro Asp Ser Arg Tyr Phe Ala Ile Ala Gln Ile Gln Gln	1185	1190	1195
Asp Gln Ala Glu Ser Tyr Ala Asp Arg Lys Gly Trp Asn Met Leu Glu	1205	1210	1215
Ala Glu Lys Trp Leu Gly Pro Asn Leu Asn	1220	1225	

&lt;210&gt; 47

&lt;211&gt; 3780

&lt;212&gt; DNA

&lt;213&gt; Agrobacterium tumefaciens

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (3777)

208

&lt;223&gt; 15887359

&lt;400&gt; 47

gtg ccc gtg ttt gac gac ctg ttt ggc cct gaa ggg gca aag cgc gac	48
Val Pro Val Phe Asp Asp Leu Phe Gly Pro Glu Gly Ala Lys Arg Asp	
1 5 10 15	
ggc gcg gaa att ttc aag gcg ttg cgc gat gcc gcc agc gaa cgc atc	96
Gly Ala Glu Ile Phe Lys Ala Leu Arg Asp Ala Ala Ser Glu Arg Ile	
20 25 30	
ctc att ctc gat ggt gcc atg ggc acg cag atc cag ggt ctc ggt ttt	144
Leu Ile Leu Asp Gly Ala Met Gly Thr Gln Ile Gln Gly Leu Gly Phe	
35 40 45	
gac gag gat cat ttt cgt ggc gac cgt ttt atc ggc tgc gcc tgt cac	192
Asp Glu Asp His Phe Arg Gly Asp Arg Phe Ile Gly Cys Ala Cys His	
50 55 60	
cag aag ggc aat aac gac ctt ctg atc ctg aca cag ccc gat gcc atc	240
Gln Lys Gly Asn Asn Asp Leu Leu Ile Leu Thr Gln Pro Asp Ala Ile	
65 70 75 80	
gag gaa atc cac tat cgc tac gcc atg gcg ggc gcg gat att ctc gaa	288
Glu Glu Ile His Tyr Arg Tyr Ala Met Ala Gly Ala Asp Ile Leu Glu	
85 90 95	
acc aac acg ttt tcc tcc acc cgc atc gcg cag gcc gat tac gag atg	336
Thr Asn Thr Phe Ser Ser Thr Arg Ile Ala Gln Ala Asp Tyr Glu Met	
100 105 110	
gag aat gcc gtc tac gat ctc aac cgc gag ggc gcg gcg atc gtg cgc	384
Glu Asn Ala Val Tyr Asp Leu Asn Arg Glu Gly Ala Ala Ile Val Arg	
115 120 125	
cgg gcg gct cag cgc gcc gag cgc gag gat ggc cgc cgc cgt ttc gtg	432
Arg Ala Ala Gln Arg Ala Glu Arg Glu Asp Gly Arg Arg Arg Phe Val	
130 135 140	
gcc ggt gcc atc ggt ccg acc aac cgc acg gcc tcg atc tcg cct gac	480
Ala Gly Ala Ile Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp	
145 150 155 160	
gtc aac aat ccc ggt tac cgc gcc gtc agt ttc gac gat ctg cgc att	528
Val Asn Asn Pro Gly Tyr Arg Ala Val Ser Phe Asp Asp Leu Arg Ile	
165 170 175	
gcc tat ggc gag cag atc gat ggc ctg atc gac ggt ggt gcc gat atc	576
Ala Tyr Gly Glu Gln Ile Asp Gly Leu Ile Asp Gly Gly Ala Asp Ile	
180 185 190	
atc ctc atc gag acg atc ttc gat acg ctg aac gcc aag gcg gcg atc	624
Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile	
195 200 205	
ttc gcc tgc gag gaa cgt ttc gag gct aag ggc atc cgc ctg ccg gtc	672
Phe Ala Cys Glu Glu Arg Phe Glu Ala Lys Gly Ile Arg Leu Pro Val	
210 215 220	
atg atc tca ggc acg atc acc gac ctt tcc ggt cgc acg ttg tcc ggc	720
Met Ile Ser Gly Thr Ile Thr Asp Leu Ser Gly Arg Thr Leu Ser Gly	
225 230 235 240	

## 209

cag acg cct tcg gcg ttc tgg aac tcg gtg cgc cac gcc aac ccc ttc Gln Thr Pro Ser Ala Phe Trp Asn Ser Val Arg His Ala Asn Pro Phe 245 250 255	768
acc atc ggc ctc aac tgc gcg ctc ggt gcg gat gcc atg cgc ccg cat Thr Ile Gly Leu Asn Cys Ala Leu Gly Ala Asp Ala Met Arg Pro His 260 265 270	816
ctg cag gaa ctg tcc gat gtg gcc gac acc ttt gtc tgc gcc tat ccg Leu Gln Glu Leu Ser Asp Val Ala Asp Thr Phe Val Cys Ala Tyr Pro 275 280 285	864
aat gcc ggc ctg ccg aac gag ttc ggc caa tat gac gaa acg ccc gag Asn Ala Gly Leu Pro Asn Glu Phe Gly Gln Tyr Asp Glu Thr Pro Glu 290 295 300	912
atg atg gcg cgc cag gtt gag ggc ttc gtt cgt gac ggt ctc gtc aac Met Met Ala Arg Gln Val Glu Gly Phe Val Arg Asp Gly Leu Val Asn 305 310 315 320	960
atc gtc ggc ggt tgc tgc ggt tcg acg ccg gaa cat atc cgg gcg att Ile Val Gly Gly Cys Cys Gly Ser Thr Pro Glu His Ile Arg Ala Ile 325 330 335	1008
gcc gaa gcc gtc aag gat tac aag ccc cgc gaa att cct gaa cac aag Ala Glu Ala Val Lys Asp Tyr Lys Pro Arg Glu Ile Pro Glu His Lys 340 345 350	1056
ccg ttc atg tcg ctt tcc ggc ctt gaa ccc ttc gtg ctg acc aag gac Pro Phe Met Ser Leu Ser Gly Leu Glu Pro Phe Val Leu Thr Lys Asp 355 360 365	1104
att ccc ttc gtc aac gtg ggc gag cgc acc aac gtc acc ggt tcg gcc Ile Pro Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala 370 375 380	1152
cgc ttc cgc aag ctc atc act gcc ggc gac tat acg gcg gcg ctg gct Arg Phe Arg Lys Leu Ile Thr Ala Gly Asp Tyr Thr Ala Ala Leu Ala 385 390 395 400	1200
gtt gcc cgc gac cag gtg gaa aac ggc gcg cag atc atc gac atc aac Val Ala Arg Asp Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile Asn 405 410 415	1248
atg gat gag ggc ctg atc gat tcg gaa aag gcg atg gtc gag ttc ctg Met Asp Glu Gly Leu Ile Asp Ser Glu Lys Ala Met Val Glu Phe Leu 420 425 430	1296
aac ctc atc gcc gcc gag cct gac att gcc cgt gtg ccc gtc atg atc Asn Leu Ile Ala Ala Glu Pro Asp Ile Ala Arg Val Pro Val Met Ile 435 440 445	1344
gac tca tcc aag ttc gag atc atc gag gcc ggc ctg aaa tgc gtg cag Asp Ser Ser Lys Phe Glu Ile Ile Glu Ala Gly Leu Lys Cys Val Gln 450 455 460	1392
ggc aaa tcg atc gtc aat tcc att tcg ctg aag gaa ggc gag gag aag Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu Lys 465 470 475 480	1440
ttt ctc cag cag gct cgg ctc gtc cac aat tac ggt gcg gcg gtt gtc	1488

210

Phe Leu Gln Gln Ala Arg Leu Val His Asn Tyr Gly Ala Ala Val Val	
485	490 495
gtc atg gcc ttt gat gag gtc ggg cag gcg gat acc tat cag cgc aag	1536
Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Tyr Gln Arg Lys	
500	505 510
gtg gaa atc tgc gcg cgc gcc tac aag ctt ctg acc gaa aag gcc ggt	1584
Val Glu Ile Cys Ala Arg Ala Tyr Lys Leu Leu Thr Glu Lys Ala Gly	
515	520 525
ctg tct ccg gaa gac atc atc ttc gac ccg aat gtg ttt gcg gta gct	1632
Leu Ser Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Phe Ala Val Ala	
530	535 540
acg ggc atc gag gag cac aat aat tac ggc gtg gac ttc atc gag gcc	1680
Thr Gly Ile Glu Glu His Asn Asn Tyr Gly Val Asp Phe Ile Glu Ala	
545	550 555 560
acc aag acc atc cgc gaa acc atg ccg ctc acg cat att tcc ggg ggc	1728
Thr Lys Thr Ile Arg Glu Thr Met Pro Leu Thr His Ile Ser Gly Gly	
565	570 575
gtt tcc aac ctg tcc ttc tcc ttc cgc ggc aat gag ccg gtg cgt gag	1776
Val Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg Glu	
580	585 590
gcg atg cat gcc gtg ttc ctc tat cac gcc att cag gtc ggc atg gat	1824
Ala Met His Ala Val Phe Leu Tyr His Ala Ile Gln Val Gly Met Asp	
595	600 605
atg ggc atc gtc aac gcc ggg cag ctt gcg gtt tac gac aat atc gat	1872
Met Gly Ile Val Asn Ala Gly Gln Leu Ala Val Tyr Asp Asn Ile Asp	
610	615 620
gcg gaa ctg cgc gag gcc tgc gaa gac gtg gtg ctg aac cgc cgc gac	1920
Ala Glu Leu Arg Glu Ala Cys Glu Asp Val Val Leu Asn Arg Arg Asp	
625	630 635 640
gat gcc acg gag cgt ctg ctc gag gtg gcg gag cgt ttc cgt ggt acg	1968
Asp Ala Thr Glu Arg Leu Leu Glu Val Ala Glu Arg Phe Arg Gly Thr	
645	650 655
ggg gaa aaa cag gcc aag gtg cag gat ctt tcc tgg cgc gag tat ccc	2016
Gly Glu Lys Gln Ala Lys Val Gln Asp Leu Ser Trp Arg Glu Tyr Pro	
660	665 670
gtt gaa aag cgg ctg gaa cat gct ctg gtc aac ggc att acc gac tat	2064
Val Glu Lys Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Asp Tyr	
675	680 685
atc gag gcc gat acg gaa gag gca cgc cag cag gcc gcc cgc ccg ctg	2112
Ile Glu Ala Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Leu	
690	695 700
cat gtc atc gaa ggg ccg ctg atg gcc ggt atg aat gtg gtg ggt gac	2160
His Val Ile Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp	
705	710 715 720
ctg ttc ggt tcc ggc aag atg ttc ctg cca cag gtg gtg aaa tcc gcc	2208
Leu Phe Gly Ser Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala	
725	730 735



## 211

cgt gtg atg aag cag gcg gtt gcc gtt ctg ctg cct tac atg gaa gag	2256
Arg Val Met Lys Gln Ala Val Ala Val Leu Leu Pro Tyr Met Glu Glu	
740 745 750	
gaa aag cgc ctg aat ggc ggt tcc gag cgc agt gcc gcc ggc aag gtg	2304
Glu Lys Arg Leu Asn Gly Gly Ser Glu Arg Ser Ala Ala Gly Lys Val	
755 760 765	
cta atg gcg acc gtg aag ggc gac gtg cac gat atc ggc aag aac atc	2352
Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile	
770 775 780	
gtc gcc gtt gtg cta gcc tgc aac aat tac gag atc att gat ctc gcc	2400
Val Gly Val Val Leu Ala Cys Asn Asn Tyr Glu Ile Ile Asp Leu Gly	
785 790 795 800	
gtg atg gtg ccg acg acg aaa atc ctc gaa acg gcg atc gcc gaa aag	2448
Val Met Val Pro Thr Thr Lys Ile Leu Glu Thr Ala Ile Ala Glu Lys	
805 810 815	
gtg gat gtg atc ggc ctc tcc ggc ctc atc acc ccg tcg ctg gat gag	2496
Val Asp Val Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu	
820 825 830	
atg gtg cat gtg gcg gcc gaa atg gag cga cag ggt ttc gac att ccg	2544
Met Val His Val Ala Ala Glu Met Glu Arg Gln Gly Phe Asp Ile Pro	
835 840 845	
ctg ctg atc ggc ggt gcg acg acc agc cgt gtg cat acg gcg gta aaa	2592
Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val Lys	
850 855 860	
atc cat ccg cgt tac gag cag ggg cag gcg atc tat gtc acc gac gcc	2640
Ile His Pro Arg Tyr Glu Gln Gly Gln Ala Ile Tyr Val Thr Asp Ala	
865 870 875 880	
tcg cgc gcg gtg ggc gtc gtt tca gcg ctc ctc tcc gaa gag cag aag	2688
Ser Arg Ala Val Gly Val Val Ser Ala Leu Leu Ser Glu Glu Gln Lys	
885 890 895	
ccc gct tat atc gac ggc atc cga gcc gaa tat gcc aag gtg gcg gaa	2736
Pro Ala Tyr Ile Asp Gly Ile Arg Ala Glu Tyr Ala Lys Val Ala Glu	
900 905 910	
gcc cat gcc cgc aat gag cgc gaa aag cag cgc ctg ccg ctt tcc cgc	2784
Ala His Ala Arg Asn Glu Arg Glu Lys Gln Arg Leu Pro Leu Ser Arg	
915 920 925	
gcc cgg gag aat gcg cac aag atc gac tgg tcg agc tac agc gtt gtc	2832
Ala Arg Glu Asn Ala His Lys Ile Asp Trp Ser Ser Tyr Ser Val Val	
930 935 940	
aag ccg cag ttc ttc ggc acc aag gtt ttt gag acc tat gat ctg gaa	2880
Lys Pro Gln Phe Phe Gly Thr Lys Val Phe Glu Thr Tyr Asp Leu Glu	
945 950 955 960	
gag ctt tcc cgt tac atc gac tgg acg ccg ttc ttc cag acc tgg gaa	2928
Glu Leu Ser Arg Tyr Ile Asp Trp Thr Pro Phe Phe Gln Thr Trp Glu	
965 970 975	
ttg aag ggc cgt ttc ccg gcg atc ctt gaa gac gaa aag cag ggc gag	2976

Leu	Lys	Gly	Arg	Phe	Pro	Ala	Ile	Leu	Glu	Asp	Glu	Lys	Gln	Gly	Glu	
980							985			990						
gcg	gcg	cgg	cag	ctt	tat	gcc	gat	gcg	cag	gcc	atg	ctt	gcg	aag	atc	3024
Ala	Ala	Arg	Gln	Leu	Tyr	Ala	Asp	Ala	Gln	Ala	Met	Leu	Ala	Lys	Ile	
995				1000				1005								
atc	gag	gaa	aag	tgg	ttc	cga	cca	cgc	gcg	gtg	atc	ggc	ttc	tgg	ccg	3072
Ile	Glu	Glu	Lys	Trp	Phe	Arg	Pro	Arg	Ala	Val	Ile	Gly	Phe	Trp	Pro	
1010				1015				1020								
gcc	aat	gcc	gtg	ggt	gac	gat	atc	agg	ctc	ttt	acg	gat	gaa	ggt	cgg	3120
Ala	Asn	Ala	Val	Gly	Asp	Asp	Ile	Arg	Leu	Phe	Thr	Asp	Glu	Gly	Arg	
1025				1030				1035				1040				
aag	gaa	gag	ttg	gcg	acg	ttc	ttc	acg	ctg	cgc	cag	cag	ctt	tcc	aag	3168
Lys	Glu	Glu	Leu	Ala	Thr	Phe	Phe	Thr	Leu	Arg	Gln	Gln	Leu	Ser	Lys	
1045				1050				1055								
cgc	gat	ggc	cgt	ccg	aac	gtg	gcg	ctg	tcc	gat	ttc	gtc	gcg	ccc	gtc	3216
Arg	Asp	Gly	Arg	Pro	Asn	Val	Ala	Leu	Ser	Asp	Phe	Val	Ala	Pro	Val	
1060				1065				1070								
gat	agc	ggc	gtt	gcc	gat	tat	gtc	ggc	ggt	ttc	gtg	gta	acg	gcg	ggt	3264
Asp	Ser	Gly	Val	Ala	Asp	Tyr	Val	Gly	Gly	Phe	Val	Val	Thr	Ala	Gly	
1075				1080				1085								
atc	gag	gaa	gtg	gcg	att	gcc	gag	cgc	ttc	gag	cgg	gcc	aat	gac	gat	3312
Ile	Glu	Glu	Val	Ala	Ile	Ala	Glu	Arg	Phe	Glu	Arg	Ala	Asn	Asp	Asp	
1090				1095				1100								
tat	tcg	tcc	atc	ctc	gtc	aag	gcg	ttg	gct	gac	cgt	ttt	gcc	gaa	gcc	3360
Tyr	Ser	Ser	Ile	Leu	Val	Lys	Ala	Leu	Ala	Asp	Arg	Phe	Ala	Glu	Ala	
1105				1110				1115				1120				
ttt	gcc	gag	cgt	atg	cat	gag	cgc	gtg	cgc	aag	gag	ttc	tgg	ggt	tat	3408
Phe	Ala	Glu	Arg	Met	His	Glu	Arg	Val	Arg	Lys	Glu	Phe	Trp	Gly	Tyr	
1125				1130				1135								
gcg	ccg	gac	gag	gct	ctt	gcc	ggt	gac	gat	ctg	ata	ggc	gaa	gcc	tat	3456
Ala	Pro	Asp	Glu	Ala	Leu	Ala	Gly	Asp	Asp	Leu	Ile	Gly	Glu	Ala	Tyr	
1140				1145				1150								
gcc	ggt	atc	cgc	ccg	gca	ccg	ggt	tat	ccg	gcc	cag	ccg	gac	cac	acc	3504
Ala	Gly	Ile	Arg	Pro	Ala	Pro	Gly	Tyr	Pro	Ala	Gln	Pro	Asp	His	Thr	
1155				1160				1165								
gaa	aag	aag	acg	ctg	ttt	gct	ctg	ctg	gac	gcc	acc	aat	gcg	gcg	ggt	3552
Glu	Lys	Lys	Thr	Leu	Phe	Ala	Leu	Leu	Asp	Ala	Thr	Asn	Ala	Ala	Gly	
1170				1175				1180								
gtg	gaa	ttg	acg	gaa	agc	tat	gcg	atg	tgg	ccc	ggc	tcg	tcg	gtt	tcg	3600
Val	Glu	Leu	Thr	Glu	Ser	Tyr	Ala	Met	Trp	Pro	Gly	Ser	Ser	Val	Ser	
1185				1190				1195				1200				
ggc	ctc	tat	atc	ggc	cat	ccc	gaa	agc	tat	tat	ttc	ggc	gtt	gcc	aag	3648
Gly	Leu	Tyr	Ile	Gly	His	Pro	Glu	Ser	Tyr	Tyr	Phe	Gly	Val	Ala	Lys	
1205				1210</												

213

gtc aca gag gtg gag cgc tgg ctc ggg ccg gtg ctc aac tac gtg ccg 3744  
 Val Thr Glu Val Glu Arg Trp Leu Gly Pro Val Leu Asn Tyr Val Pro  
 1235 1240 1245

acc aac ggc gag gag aaa atc gac agc gct gcg tga 3780  
 Thr Asn Gly Glu Glu Lys Ile Asp Ser Ala Ala  
 1250 1255

<210> 48  
 <211> 1259  
 <212> PRT  
 <213> Agrobacterium tumefaciens

<400> 48  
 Val Pro Val Phe Asp Asp Leu Phe Gly Pro Glu Gly Ala Lys Arg Asp  
 1 5 10 15  
 Gly Ala Glu Ile Phe Lys Ala Leu Arg Asp Ala Ala Ser Glu Arg Ile  
 20 25 30  
 Leu Ile Leu Asp Gly Ala Met Gly Thr Gln Ile Gln Gly Leu Gly Phe  
 35 40 45  
 Asp Glu Asp His Phe Arg Gly Asp Arg Phe Ile Gly Cys Ala Cys His  
 50 55 60  
 Gln Lys Gly Asn Asn Asp Leu Leu Ile Leu Thr Gln Pro Asp Ala Ile  
 65 70 75 80  
 Glu Glu Ile His Tyr Arg Tyr Ala Met Ala Gly Ala Asp Ile Leu Glu  
 85 90 95  
 Thr Asn Thr Phe Ser Ser Thr Arg Ile Ala Gln Ala Asp Tyr Glu Met  
 100 105 110  
 Glu Asn Ala Val Tyr Asp Leu Asn Arg Glu Gly Ala Ala Ile Val Arg  
 115 120 125  
 Arg Ala Ala Gln Arg Ala Glu Arg Glu Asp Gly Arg Arg Arg Phe Val  
 130 135 140  
 Ala Gly Ala Ile Gly Pro Thr Asn Arg Thr Ala Ser Ile Ser Pro Asp  
 145 150 155 160  
 Val Asn Asn Pro Gly Tyr Arg Ala Val Ser Phe Asp Asp Leu Arg Ile  
 165 170 175  
 Ala Tyr Gly Glu Gln Ile Asp Gly Leu Ile Asp Gly Gly Ala Asp Ile  
 180 185 190  
 Ile Leu Ile Glu Thr Ile Phe Asp Thr Leu Asn Ala Lys Ala Ala Ile  
 195 200 205  
 Phe Ala Cys Glu Glu Arg Phe Glu Ala Lys Gly Ile Arg Leu Pro Val  
 210 215 220  
 Met Ile Ser Gly Thr Ile Thr Asp Leu Ser Gly Arg Thr Leu Ser Gly  
 225 230 235 240  
 Gln Thr Pro Ser Ala Phe Trp Asn Ser Val Arg His Ala Asn Pro Phe

214

245	250	255
Thr Ile Gly Leu Asn Cys Ala Leu Gly Ala Asp Ala Met Arg Pro His 260	265	270
Leu Gln Glu Leu Ser Asp Val Ala Asp Thr Phe Val Cys Ala Tyr Pro 275	280	285
Asn Ala Gly Leu Pro Asn Glu Phe Gly Gln Tyr Asp Glu Thr Pro Glu 290	295	300
Met Met Ala Arg Gln Val Glu Gly Phe Val Arg Asp Gly Leu Val Asn 305	310	315
Ile Val Gly Gly Cys Cys Gly Ser Thr Pro Glu His Ile Arg Ala Ile 325	330	335
Ala Glu Ala Val Lys Asp Tyr Lys Pro Arg Glu Ile Pro Glu His Lys 340	345	350
Pro Phe Met Ser Leu Ser Gly Leu Glu Pro Phe Val Leu Thr Lys Asp 355	360	365
Ile Pro Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser Ala 370	375	380
Arg Phe Arg Lys Leu Ile Thr Ala Gly Asp Tyr Thr Ala Ala Leu Ala 385	390	395
Val Ala Arg Asp Gln Val Glu Asn Gly Ala Gln Ile Ile Asp Ile Asn 405	410	415
Met Asp Glu Gly Leu Ile Asp Ser Glu Lys Ala Met Val Glu Phe Leu 420	425	430
Asn Leu Ile Ala Ala Glu Pro Asp Ile Ala Arg Val Pro Val Met Ile 435	440	445
Asp Ser Ser Lys Phe Glu Ile Ile Glu Ala Gly Leu Lys Cys Val Gln 450	455	460
Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu Lys 465	470	475
Phe Leu Gln Gln Ala Arg Leu Val His Asn Tyr Gly Ala Ala Val Val 485	490	495
Val Met Ala Phe Asp Glu Val Gly Gln Ala Asp Thr Tyr Gln Arg Lys 500	505	510
Val Glu Ile Cys Ala Arg Ala Tyr Lys Leu Leu Thr Glu Lys Ala Gly 515	520	525
Leu Ser Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Phe Ala Val Ala 530	535	540
Thr Gly Ile Glu Glu His Asn Asn Tyr Gly Val Asp Phe Ile Glu Ala 545	550	555
Thr Lys Thr Ile Arg Glu Thr Met Pro Leu Thr His Ile Ser Gly Gly 565	570	575

215

Val Ser Asn Leu Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg Glu  
 580 585 590  
 Ala Met His Ala Val Phe Leu Tyr His Ala Ile Gln Val Gly Met Asp  
 595 600 605  
 Met Gly Ile Val Asn Ala Gly Gln Leu Ala Val Tyr Asp Asn Ile Asp  
 610 615 620  
 Ala Glu Leu Arg Glu Ala Cys Glu Asp Val Val Leu Asn Arg Arg Asp  
 625 630 635 640  
 Asp Ala Thr Glu Arg Leu Leu Glu Val Ala Glu Arg Phe Arg Gly Thr  
 645 650 655  
 Gly Glu Lys Gln Ala Lys Val Gln Asp Leu Ser Trp Arg Glu Tyr Pro  
 660 665 670  
 Val Glu Lys Arg Leu Glu His Ala Leu Val Asn Gly Ile Thr Asp Tyr  
 675 680 685  
 Ile Glu Ala Asp Thr Glu Glu Ala Arg Gln Gln Ala Ala Arg Pro Leu  
 690 695 700  
 His Val Ile Glu Gly Pro Leu Met Ala Gly Met Asn Val Val Gly Asp  
 705 710 715 720  
 Leu Phe Gly Ser Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala  
 725 730 735  
 Arg Val Met Lys Gln Ala Val Ala Val Leu Leu Pro Tyr Met Glu Glu  
 740 745 750  
 Glu Lys Arg Leu Asn Gly Gly Ser Glu Arg Ser Ala Ala Gly Lys Val  
 755 760 765  
 Leu Met Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile  
 770 775 780  
 Val Gly Val Val Leu Ala Cys Asn Asn Tyr Glu Ile Ile Asp Leu Gly  
 785 790 795 800  
 Val Met Val Pro Thr Thr Lys Ile Leu Glu Thr Ala Ile Ala Glu Lys  
 805 810 815  
 Val Asp Val Ile Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Asp Glu  
 820 825 830  
 Met Val His Val Ala Ala Glu Met Glu Arg Gln Gly Phe Asp Ile Pro  
 835 840 845  
 Leu Leu Ile Gly Gly Ala Thr Thr Ser Arg Val His Thr Ala Val Lys  
 850 855 860  
 Ile His Pro Arg Tyr Glu Gln Gly Gln Ala Ile Tyr Val Thr Asp Ala  
 865 870 875 880  
 Ser Arg Ala Val Gly Val Val Ser Ala Leu Leu Ser Glu Glu Gln Lys  
 885 890 895  
 Pro Ala Tyr Ile Asp Gly Ile Arg Ala Glu Tyr Ala Lys Val Ala Glu  
 900 905 910

216

Ala His Ala Arg Asn Glu Arg Glu Lys Gln Arg Leu Pro Leu Ser Arg  
 915 920 925

Ala Arg Glu Asn Ala His Lys Ile Asp Trp Ser Ser Tyr Ser Val Val  
 930 935 940

Lys Pro Gln Phe Phe Gly Thr Lys Val Phe Glu Thr Tyr Asp Leu Glu  
 945 950 955 960

Glu Leu Ser Arg Tyr Ile Asp Trp Thr Pro Phe Phe Gln Thr Trp Glu  
 965 970 975

Leu Lys Gly Arg Phe Pro Ala Ile Leu Glu Asp Glu Lys Gln Gly Glu  
 980 985 990

Ala Ala Arg Gln Leu Tyr Ala Asp Ala Gln Ala Met Leu Ala Lys Ile  
 995 1000 1005

Ile Glu Glu Lys Trp Phe Arg Pro Arg Ala Val Ile Gly Phe Trp Pro  
 1010 1015 1020

Ala Asn Ala Val Gly Asp Asp Ile Arg Leu Phe Thr Asp Glu Gly Arg  
 1025 1030 1035 1040

Lys Glu Glu Leu Ala Thr Phe Phe Thr Leu Arg Gln Gln Leu Ser Lys  
 1045 1050 1055

Arg Asp Gly Arg Pro Asn Val Ala Leu Ser Asp Phe Val Ala Pro Val  
 1060 1065 1070

Asp Ser Gly Val Ala Asp Tyr Val Gly Gly Phe Val Val Thr Ala Gly  
 1075 1080 1085

Ile Glu Glu Val Ala Ile Ala Glu Arg Phe Glu Arg Ala Asn Asp Asp  
 1090 1095 1100

Tyr Ser Ser Ile Leu Val Lys Ala Leu Ala Asp Arg Phe Ala Glu Ala  
 1105 1110 1115 1120

Phe Ala Glu Arg Met His Glu Arg Val Arg Lys Glu Phe Trp Gly Tyr  
 1125 1130 1135

Ala Pro Asp Glu Ala Leu Ala Gly Asp Asp Leu Ile Gly Glu Ala Tyr  
 1140 1145 1150

Ala Gly Ile Arg Pro Ala Pro Gly Tyr Pro Ala Gln Pro Asp His Thr  
 1155 1160 1165

Glu Lys Lys Thr Leu Phe Ala Leu Leu Asp Ala Thr Asn Ala Ala Gly  
 1170 1175 1180

Val Glu Leu Thr Glu Ser Tyr Ala Met Trp Pro Gly Ser Ser Val Ser  
 1185 1190 1195 1200

Gly Leu Tyr Ile Gly His Pro Glu Ser Tyr Tyr Phe Gly Val Ala Lys  
 1205 1210 1215

Val Glu Arg Asp Gln Val Leu Asp Tyr Ala Arg Arg Lys Asp Met Pro  
 1220 1225 1230

Val Thr Glu Val Glu Arg Trp Leu Gly Pro Val Leu Asn Tyr Val Pro

217

1235

1240

1245

Thr Asn Gly Glu Glu Lys Ile Asp Ser Ala Ala  
1250 1255

&lt;210&gt; 49

&lt;211&gt; 2718

&lt;212&gt; DNA

&lt;213&gt; Ralstonia solanacearum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(2715)

&lt;223&gt; RSOL\_GMI1000

&lt;400&gt; 49

```

atg acc gac cac ctc atg cgc ctc tcc ggc ctc gaa ccg ttc aac atc 48
Met Thr Asp His Leu Met Arg Leu Ser Gly Leu Glu Pro Phe Asn Ile
  1             5             10             15

ggc gag gac acg ctg ttc gtc aac gtc ggc gaa cgc acc aac gtc acc 96
Gly Glu Asp Thr Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr
          20             25             30

gga tcc aag gcg ttc gcg cgc atg atc ctc aac agc cag ttc gac gag 144
Gly Ser Lys Ala Phe Ala Arg Met Ile Leu Asn Ser Gln Phe Asp Glu
          35             40             45

gcg ctc gcc gtg gca cgc cag cag gtc gag aac ggc gcg cag gtc atc 192
Ala Leu Ala Val Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile
          50             55             60

gac atc aac atg gac gag gcc atg ctc gac tcc aag gcg gcg atg gtg 240
Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Lys Ala Ala Met Val
          65             70             75             80

cgc ttc ctg aac ctg atc gcc tcg gag ccg gac atc gcg cgc gtg ccg 288
Arg Phe Leu Asn Leu Ile Ala Ser Glu Pro Asp Ile Ala Arg Val Pro
          85             90             95

atc atg atc gac tcg tcc aag tgg gag gtg atc gag gcc ggc ctg aag 336
Ile Met Ile Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys
          100             105             110

tgc gtg cag ggc aag gcc atc gtc aac tcg atc tcg ctc aag gaa ggc 384
Cys Val Gln Gly Lys Ala Ile Val Asn Ser Ile Ser Leu Lys Glu Gly
          115             120             125

gag gaa cag ttc gcc cac cac gcc aag ctg atc aag cgc tac ggc gcc 432
Glu Glu Gln Phe Ala His His Ala Lys Leu Ile Lys Arg Tyr Gly Ala
          130             135             140

gcc gcc gtg gtg atg gcc ttc gac gag cag ggc cag gcc gac acg ttc 480
Ala Ala Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Phe
          145             150             155             160

gcg cgc aag acc gag atc tgc aag cgc agc tat gac ttc ctc gtg aac 528
Ala Arg Lys Thr Glu Ile Cys Lys Arg Ser Tyr Asp Phe Leu Val Asn
          165             170             175

cag gtc ggc ttt gcg ccg gaa gac atc atc ttc gat ccg aac atc ttc 576

```

218

Gln Val Gly Phe Ala Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe	
180 185 190	
gcg gtc gcc acc ggc atc gag gag cac aac aac tac gcc gtc gac ttc	624
Ala Val Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Val Asp Phe	
195 200 205	
atc gag gcc acg cgc tgg atc aag cag aaa ttg ccg cac gcc aag gtg	672
Ile Glu Ala Thr Arg Trp Ile Lys Gln Lys Leu Pro His Ala Lys Val	
210 215 220	
agc ggc ggc gtg tcg aac gtc tcg ttc tcg ttc cgc ggc aac gac gtg	720
Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Val	
225 230 235 240	
gtg cgc gag gcc atc cac acc gtg ttc ctg tac cac gcc atc ggt gcg	768
Val Arg Glu Ala Ile His Thr Val Phe Leu Tyr His Ala Ile Gly Ala	
245 250 255	
ggc atg gac atg ggc atc gtc aac gcg ggc cag ttg ggc gtg tac gag	816
Gly Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Gly Val Tyr Glu	
260 265 270	
aac ctc gcc ccc gaa ctg cgc gag cgc gtg gaa gac gtg gtg ctc aac	864
Asn Leu Ala Pro Glu Leu Arg Glu Arg Val Glu Asp Val Val Leu Asn	
275 280 285	
cgc cgc ccg gat gcg acc gac cgc ctg ctg gaa att gcc gac cgc tac	912
Arg Arg Pro Asp Ala Thr Asp Arg Leu Leu Glu Ile Ala Asp Arg Tyr	
290 295 300	
aag ggc ggc ggc gcc aag cgc gag gag aac ctc gcc tgg cgc cag gag	960
Lys Gly Gly Gly Ala Lys Arg Glu Glu Asn Leu Ala Trp Arg Gln Glu	
305 310 315 320	
ccg gtg gaa aag cgc ctg gcc cac gcg ctc gtg cac ggc atc acc gac	1008
Pro Val Glu Lys Arg Leu Ala His Ala Leu Val His Gly Ile Thr Asp	
325 330 335	
tac gtg gtc gaa gac acc gag gaa gtt cgc cag aag atc ttt gcc gcc	1056
Tyr Val Val Glu Asp Thr Glu Glu Val Arg Gln Lys Ile Phe Ala Ala	
340 345 350	
ggc ggc cgc ccg atc cag gtg atc gag ggc ccg ctg atg gac ggc atg	1104
Gly Gly Arg Pro Ile Gln Val Ile Glu Gly Pro Leu Met Asp Gly Met	
355 360 365	
aac atc gtc ggc gat ctg ttc ggc gcg ggc aag atg ttc ctg ccg cag	1152
Asn Ile Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln	
370 375 380	
gtg gtg aaa tcc gcc cgc gtg atg aag cag gcg gtg gcc cac ctg atc	1200
Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Ile	
385 390 395 400	
ccg ttc atc gag gaa gag aag cgg cag atc gcg gcc gcc ggc ggc gac	1248
Pro Phe Ile Glu Glu Lys Arg Gln Ile Ala Ala Ala Gly Gly Asp	
405 410 415	
gtg cgc tcg cgc ggc aag atc gtc atc gcc acc gtg aag ggc gac gtg	1296
Val Arg Ser Arg Gly Lys Ile Val Ile Ala Thr Val Lys Gly Asp Val	
420 425 430	



219

cac gac atc ggc aag aac atc gtc acc gtc gtg ctc cag tgc aac aac	1344
His Asp Ile Gly Lys Asn Ile Val Thr Val Val Leu Gln Cys Asn Asn	
435 440 445	
ttc gaa gtc gtg aac atg ggc gtg atg gtc ccg tgc aac gag atc ctg	1392
Phe Glu Val Val Asn Met Gly Val Met Val Pro Cys Asn Glu Ile Leu	
450 455 460	
gcc aag gcg aag gtc gag ggc gcg gac atc atc ggc ctg tcg ggc ctg	1440
Ala Lys Ala Lys Val Glu Gly Ala Asp Ile Ile Gly Leu Ser Gly Leu	
465 470 475 480	
atc aca ccg tcg ctg gaa gag atg gcc tac gtg gcc tcc gag atg cag	1488
Ile Thr Pro Ser Leu Glu Glu Met Ala Tyr Val Ala Ser Glu Met Gln	
485 490 495	
cgc gac gag tac ttc cgc gtg aag aag atc ccg ctg ctg atc ggt ggc	1536
Arg Asp Glu Tyr Phe Arg Val Lys Lys Ile Pro Leu Leu Ile Gly Gly	
500 505 510	
gcg acc acg agc cgc gtg cac acc gcc gtg aag atc gcg ccc aat tac	1584
Ala Thr Thr Ser Arg Val His Thr Ala Val Lys Ile Ala Pro Asn Tyr	
515 520 525	
gaa ggc ccg gtc gtg tac gtg ccc gac gcc tcg cgc tcg gtg agc gtg	1632
Glu Gly Pro Val Val Tyr Val Pro Asp Ala Ser Arg Ser Val Ser Val	
530 535 540	
gcc tcc agc ctg ctg tcc gac gag gcc gcc gcg cgc tac atc gaa gag	1680
Ala Ser Ser Leu Leu Ser Asp Glu Ala Ala Ala Arg Tyr Ile Glu Glu	
545 550 555 560	
ctg cac gcc gac tac gac cgc atc cgc acc cag cac gcc agc aag aaa	1728
Leu His Ala Asp Tyr Asp Arg Ile Arg Thr Gln His Ala Ser Lys Lys	
565 570 575	
gcc atg ccg atg gtg tcg ctg gcc gcc gcg cgc gcc aac aag acc cgg	1776
Ala Met Pro Met Val Ser Leu Ala Ala Ala Arg Ala Asn Lys Thr Arg	
580 585 590	
atc gac tgg tcg aac tac acg ccg ccc aag ccc aag ttc gtc ggc cgc	1824
Ile Asp Trp Ser Asn Tyr Thr Pro Pro Lys Pro Lys Phe Val Gly Arg	
595 600 605	
cgc gtg ttc cgc aac tac gac ctg aac gag ctc gcg cag tac atc gac	1872
Arg Val Phe Arg Asn Tyr Asp Leu Asn Glu Leu Ala Gln Tyr Ile Asp	
610 615 620	
tgg ggc ccg ttc ttc cag acg tgg gac ctg gcc ggc aaa ttc ccc gac	1920
Trp Gly Pro Phe Phe Gln Thr Trp Asp Leu Ala Gly Lys Phe Pro Asp	
625 630 635 640	
atc ctc aac gac gcg atc gtc ggc gaa tcg gcc cgc cgc gtg ttc tcc	1968
Ile Leu Asn Asp Ala Ile Val Gly Glu Ser Ala Arg Arg Val Phe Ser	
645 650 655	
gac ggc aag agc atg ctc gcg cgc ctg atc gcc gga cgc tgg ctg acg	2016
Asp Gly Lys Ser Met Leu Ala Arg Leu Ile Ala Gly Arg Trp Leu Thr	
660 665 670	
gcc aac ggc gtg atc gcg ctg ctg ccg gcc aac acc gtc aac gac gac	2064

220

Ala Asn Gly Val Ile Ala Leu Leu Pro Ala Asn Thr Val Asn Asp Asp	
675 680 685	
gac atc gag atc tac acc gac gag acc cgc tcg gaa gtc gcc ctc acc	2112
Asp Ile Glu Ile Tyr Thr Asp Glu Thr Arg Ser Glu Val Ala Leu Thr	
690 695 700	
tgg cgc aac atc cgc cag cag agc gag cgc ccg atc atc gac ggc gtg	2160
Trp Arg Asn Ile Arg Gln Gln Ser Glu Arg Pro Ile Ile Asp Gly Val	
705 710 715 720	
atg cgc ccg aac cgc tgc ctg gcg gac ttc atc gcc ccc aag gac acc	2208
Met Arg Pro Asn Arg Cys Leu Ala Asp Phe Ile Ala Pro Lys Asp Thr	
725 730 735	
ggc atc gcc gat tac atc ggc ctc ttc gcg gtg acg ggc ggc atc ggg	2256
Gly Ile Ala Asp Tyr Ile Gly Leu Phe Ala Val Thr Gly Gly Ile Gly	
740 745 750	
atc gac aag cgc gaa gcc gcc ttc gaa gcc gac cac gac gac tac agc	2304
Ile Asp Lys Arg Glu Ala Ala Phe Glu Ala Asp His Asp Asp Tyr Ser	
755 760 765	
gcg atc atg ctc aag gcc ctg gcc gac cgc ttc gcc gaa gcc ttc gcc	2352
Ala Ile Met Leu Lys Ala Leu Ala Asp Arg Phe Ala Glu Ala Phe Ala	
770 775 780	
gag tgc ctg cac gcc cgt gtg cgc cgc gac ctg tgg ggc tac gcg cag	2400
Glu Cys Leu His Ala Arg Val Arg Arg Asp Leu Trp Gly Tyr Ala Gln	
785 790 795 800	
gac gaa acg ctc gac aac gac gcg ctg atc cgc gag gaa tac cgc ggc	2448
Asp Glu Thr Leu Asp Asn Asp Ala Leu Ile Arg Glu Glu Tyr Arg Gly	
805 810 815	
atc cgc ccg gcg ccc ggc tac ccg gcc tgc ccg gag cac acc gtc aag	2496
Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Val Lys	
820 825 830	
cgc gac ctg ttc cgc gtg ctc gac gcg cag gag atc ggc atg aac ctg	2544
Arg Asp Leu Phe Arg Val Leu Asp Ala Gln Glu Ile Gly Met Asn Leu	
835 840 845	
acc gag gcg ctg gcg atg aca ccg gcc gcg tcg gtc tcg ggc ttc cag	2592
Thr Glu Ala Leu Ala Met Thr Pro Ala Ala Ser Val Ser Gly Phe Gln	
850 855 860	
ctg tcg cac ccg gac agc acg tac ttc acg atc ggc aag atc ggc cag	2640
Leu Ser His Pro Asp Ser Thr Tyr Phe Thr Ile Gly Lys Ile Gly Gln	
865 870 875 880	
gac cag gtg gac gac atg gcc gcg cgc agc ggg gaa gac cgc cgc aat	2688
Asp Gln Val Asp Asp Met Ala Ala Arg Ser Gly Glu Asp Arg Arg Asn	
885 890 895	
gtg gag cgc gcc ctg gca ccc aac ctg taa	2718
Val Glu Arg Ala Leu Ala Pro Asn Leu	
900 905	

&lt;210&gt; 50

&lt;211&gt; 905

221

&lt;212&gt; PRT

<213> *Ralstonia solanacearum*

&lt;400&gt; 50

Met Thr Asp His Leu Met Arg Leu Ser Gly Leu Glu Pro Phe Asn Ile  
 1 5 10 15

Gly Glu Asp Thr Leu Phe Val Asn Val Gly Glu Arg Thr Asn Val Thr  
 20 25 30

Gly Ser Lys Ala Phe Ala Arg Met Ile Leu Asn Ser Gln Phe Asp Glu  
 35 40 45

Ala Leu Ala Val Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile  
 50 55 60

Asp Ile Asn Met Asp Glu Ala Met Leu Asp Ser Lys Ala Ala Met Val  
 65 70 75 80

Arg Phe Leu Asn Leu Ile Ala Ser Glu Pro Asp Ile Ala Arg Val Pro  
 85 90 95

Ile Met Ile Asp Ser Ser Lys Trp Glu Val Ile Glu Ala Gly Leu Lys  
 100 105 110

Cys Val Gln Gly Lys Ala Ile Val Asn Ser Ile Ser Leu Lys Glu Gly  
 115 120 125

Glu Glu Gln Phe Ala His His Ala Lys Leu Ile Lys Arg Tyr Gly Ala  
 130 135 140

Ala Ala Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Thr Phe  
 145 150 155 160

Ala Arg Lys Thr Glu Ile Cys Lys Arg Ser Tyr Asp Phe Leu Val Asn  
 165 170 175

Gln Val Gly Phe Ala Pro Glu Asp Ile Ile Phe Asp Pro Asn Ile Phe  
 180 185 190

Ala Val Ala Thr Gly Ile Glu Glu His Asn Asn Tyr Ala Val Asp Phe  
 195 200 205

Ile Glu Ala Thr Arg Trp Ile Lys Gln Lys Leu Pro His Ala Lys Val  
 210 215 220

Ser Gly Gly Val Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Asp Val  
 225 230 235 240

Val Arg Glu Ala Ile His Thr Val Phe Leu Tyr His Ala Ile Gly Ala  
 245 250 255

Gly Met Asp Met Gly Ile Val Asn Ala Gly Gln Leu Gly Val Tyr Glu  
 260 265 270

Asn Leu Ala Pro Glu Leu Arg Glu Arg Val Glu Asp Val Val Leu Asn  
 275 280 285

Arg Arg Pro Asp Ala Thr Asp Arg Leu Leu Glu Ile Ala Asp Arg Tyr  
 290 295 300

Lys Gly Gly Gly Ala Lys Arg Glu Glu Asn Leu Ala Trp Arg Gln Glu

222

305	310	315	320
Pro Val Glu Lys Arg Leu Ala His Ala Leu Val His Gly Ile Thr Asp	325	330	335
Tyr Val Val Glu Asp Thr Glu Glu Val Arg Gln Lys Ile Phe Ala Ala	340	345	350
Gly Gly Arg Pro Ile Gln Val Ile Glu Gly Pro Leu Met Asp Gly Met	355	360	365
Asn Ile Val Gly Asp Leu Phe Gly Ala Gly Lys Met Phe Leu Pro Gln	370	375	380
Val Val Lys Ser Ala Arg Val Met Lys Gln Ala Val Ala His Leu Ile	385	390	395
Pro Phe Ile Glu Glu Glu Lys Arg Gln Ile Ala Ala Ala Gly Gly Asp	405	410	415
Val Arg Ser Arg Gly Lys Ile Val Ile Ala Thr Val Lys Gly Asp Val	420	425	430
His Asp Ile Gly Lys Asn Ile Val Thr Val Val Leu Gln Cys Asn Asn	435	440	445
Phe Glu Val Val Asn Met Gly Val Met Val Pro Cys Asn Glu Ile Leu	450	455	460
Ala Lys Ala Lys Val Glu Gly Ala Asp Ile Ile Gly Leu Ser Gly Leu	465	470	475
Ile Thr Pro Ser Leu Glu Glu Met Ala Tyr Val Ala Ser Glu Met Gln	485	490	495
Arg Asp Glu Tyr Phe Arg Val Lys Lys Ile Pro Leu Leu Ile Gly Gly	500	505	510
Ala Thr Thr Ser Arg Val His Thr Ala Val Lys Ile Ala Pro Asn Tyr	515	520	525
Glu Gly Pro Val Val Tyr Val Pro Asp Ala Ser Arg Ser Val Ser Val	530	535	540
Ala Ser Ser Leu Leu Ser Asp Glu Ala Ala Ala Arg Tyr Ile Glu Glu	545	550	555
Leu His Ala Asp Tyr Asp Arg Ile Arg Thr Gln His Ala Ser Lys Lys	565	570	575
Ala Met Pro Met Val Ser Leu Ala Ala Ala Arg Ala Asn Lys Thr Arg	580	585	590
Ile Asp Trp Ser Asn Tyr Thr Pro Pro Lys Pro Lys Phe Val Gly Arg	595	600	605
Arg Val Phe Arg Asn Tyr Asp Leu Asn Glu Leu Ala Gln Tyr Ile Asp	610	615	620
Trp Gly Pro Phe Phe Gln Thr Trp Asp Leu Ala Gly Lys Phe Pro Asp	625	630	635
			640

223

Ile Leu Asn Asp Ala Ile Val Gly Glu Ser Ala Arg Arg Val Phe Ser  
645 650 655

Asp Gly Lys Ser Met Leu Ala Arg Leu Ile Ala Gly Arg Trp Leu Thr  
660 665 670

Ala Asn Gly Val Ile Ala Leu Leu Pro Ala Asn Thr Val Asn Asp Asp  
675 680 685

Asp Ile Glu Ile Tyr Thr Asp Glu Thr Arg Ser Glu Val Ala Leu Thr  
690 695 700

Trp Arg Asn Ile Arg Gln Gln Ser Glu Arg Pro Ile Ile Asp Gly Val  
705 710 715 720

Met Arg Pro Asn Arg Cys Leu Ala Asp Phe Ile Ala Pro Lys Asp Thr  
725 730 735

Gly Ile Ala Asp Tyr Ile Gly Leu Phe Ala Val Thr Gly Gly Ile Gly  
740 745 750

Ile Asp Lys Arg Glu Ala Ala Phe Glu Ala Asp His Asp Asp Tyr Ser  
755 760 765

Ala Ile Met Leu Lys Ala Leu Ala Asp Arg Phe Ala Glu Ala Phe Ala  
770 775 780

Glu Cys Leu His Ala Arg Val Arg Arg Asp Leu Trp Gly Tyr Ala Gln  
785 790 795 800

Asp Glu Thr Leu Asp Asn Asp Ala Leu Ile Arg Glu Glu Tyr Arg Gly  
805 810 815

Ile Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Glu His Thr Val Lys  
820 825 830

Arg Asp Leu Phe Arg Val Leu Asp Ala Gln Glu Ile Gly Met Asn Leu  
835 840 845

Thr Glu Ala Leu Ala Met Thr Pro Ala Ala Ser Val Ser Gly Phe Gln  
850 855 860

Leu Ser His Pro Asp Ser Thr Tyr Phe Thr Ile Gly Lys Ile Gly Gln  
865 870 875 880

Asp Gln Val Asp Asp Met Ala Ala Arg Ser Gly Glu Asp Arg Arg Asn  
885 890 895

Val Glu Arg Ala Leu Ala Pro Asn Leu  
900 905

&lt;210&gt; 51

&lt;211&gt; 3645

&lt;212&gt; DNA

&lt;213&gt; Chlorobium tepidum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (3642)

&lt;223&gt; RCL00420

224

&lt;400&gt; 51

gtg ctc gac ggg gcc atg ggc acc atg atc cag agg cat ggc ctc gac	48
Val Leu Asp Gly Ala Met Gly Thr Met Ile Gln Arg His Gly Leu Asp	
1 5 10 15	
gaa cag gac tac cgg ggc gag cgt ttc gct tcg cat gac cat ccg ctg	96
Glu Gln Asp Tyr Arg Gly Glu Arg Phe Ala Ser His Asp His Pro Leu	
20 25 30	
aag ggc aac aac gac ctt ctt gtc atc acc cgg ccc gac atc atc cgt	144
Lys Gly Asn Asn Asp Leu Leu Val Ile Thr Arg Pro Asp Ile Ile Arg	
35 40 45	
tcg atc cac tgc gac ttc ctc gac gcg ggt gcg gac atc atc gag acc	192
Ser Ile His Cys Asp Phe Leu Asp Ala Gly Ala Asp Ile Ile Glu Thr	
50 55 60	
tgc acc ttc aac gcc aac ccg atc tcg cag tcg gac tac cag ttg cag	240
Cys Thr Phe Asn Ala Asn Pro Ile Ser Gln Ser Asp Tyr Gln Leu Gln	
65 70 75 80	
gac ttg acc cgc gag ctg aac gtg gcg gcg gca aag ata gcc cgc tcg	288
Asp Leu Thr Arg Glu Leu Asn Val Ala Ala Ala Lys Ile Ala Arg Ser	
85 90 95	
gca gcg gac gag ttc acc gca aag act ccc gac aag ccg cgt ttc gtg	336
Ala Ala Asp Glu Phe Thr Ala Lys Thr Pro Asp Lys Pro Arg Phe Val	
100 105 110	
gcc ggt tcc atc gga ccg acc aac aag acg ctc tcg ctc tcg ccg gac	384
Ala Gly Ser Ile Gly Pro Thr Asn Lys Thr Leu Ser Leu Ser Pro Asp	
115 120 125	
gtg aac aac ccc ggc ttc cgc gcc gtc acc ttc cag gag atg gtc gat	432
Val Asn Asn Pro Gly Phe Arg Ala Val Thr Phe Gln Glu Met Val Asp	
130 135 140	
aac tac act gcc cag ctc gaa ggc ttg cac gag ggc ggt gtc gat ctc	480
Asn Tyr Thr Ala Gln Leu Glu Gly Leu His Glu Gly Gly Val Asp Leu	
145 150 155 160	
ttg ctc gtc gag acg gtg ttc gac aca ctg aac tgc aag gcg gcg ctc	528
Leu Leu Val Glu Thr Val Phe Asp Thr Leu Asn Cys Lys Ala Ala Leu	
165 170 175	
tac gct atc gag gag tac gcg gtg aaa acc ggc tgg cag gtg ccc gtg	576
Tyr Ala Ile Glu Glu Tyr Ala Val Lys Thr Gly Trp Gln Val Pro Val	
180 185 190	
atg gtc tcc ggc acg gtg gtg gac gcg agc ggc cgc acc ctc tcc ggc	624
Met Val Ser Gly Thr Val Val Asp Ala Ser Gly Arg Thr Leu Ser Gly	
195 200 205	
caa acc acc gag gcg ttc tgg att tcg att tcg cac atg ccg agt ctg	672
Gln Thr Thr Glu Ala Phe Trp Ile Ser Ile Ser His Met Pro Ser Leu	
210 215 220	
ctc tcg gtc ggc ctg aac tgc gca ctc ggc tcc aag cag atg cgc ccc	720
Leu Ser Val Gly Leu Asn Cys Ala Leu Gly Ser Lys Gln Met Arg Pro	
225 230 235 240	
ttc atc gag gcg ctc tcg aac atc gcc gaa agc tac gtc agc gtc tat	768

225

Phe Ile Glu Ala Leu Ser Asn Ile Ala Glu Ser Tyr Val Ser Val Tyr	
245 250 255	
ccc aac gcg ggc ctg ccg aat gag ttc ggc gag tac gac gac tcc ccc	816
Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp Asp Ser Pro	
260 265 270	
gag tac atg gcc gcg cag atc gcg ggc ttc gcc gaa tca ggc ttc gtg	864
Glu Tyr Met Ala Ala Gln Ile Ala Gly Phe Ala Glu Ser Gly Phe Val	
275 280 285	
aac atc gtc ggc ggc tgc tgc ggc acc acg ccg acg cac atc cgc gcc	912
Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Thr His Ile Arg Ala	
290 295 300	
att gcc gaa gcg gtc aag act ctc ccg ccg aga aag cgc ccc gcc aac	960
Ile Ala Glu Ala Val Lys Thr Leu Pro Pro Arg Lys Arg Pro Ala Asn	
305 310 315 320	
aag cac gtg ctg agg ctc tcc ggc ctc gaa ccg ctc gtg gtt gac gaa	1008
Lys His Val Leu Arg Leu Ser Gly Leu Glu Pro Leu Val Val Asp Glu	
325 330 335	
acc acc ggc ttc atc aac gtc ggc gag cgc acc aac gtc acc ggt tcg	1056
Thr Thr Gly Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser	
340 345 350	
cgc aag ttc gcc cgc ctc atc aag gag gcc aat tac gac gaa gcg ctc	1104
Arg Lys Phe Ala Arg Leu Ile Lys Glu Ala Asn Tyr Asp Glu Ala Leu	
355 360 365	
tcc att gcc cgc cag cag gtc gag aac ggc gcg cag gtg atc gac gtg	1152
Ser Ile Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile Asp Val	
370 375 380	
aac ctc gac gaa gga atg ctc gac tcc gaa aag gtg atc gtc gaa ttc	1200
Asn Leu Asp Glu Gly Met Leu Asp Ser Glu Lys Val Ile Val Glu Phe	
385 390 395 400	
ctg aac ctc atc gcc tcc gag cct gag atc gcc aag gtg ccg gtg atg	1248
Leu Asn Leu Ile Ala Ser Glu Pro Glu Ile Ala Lys Val Pro Val Met	
405 410 415	
atc gac tcg tcg aaa tgg tcg gtc atc gaa aac ggc ctg cgc tgc acc	1296
Ile Asp Ser Ser Lys Trp Ser Val Ile Glu Asn Gly Leu Arg Cys Thr	
420 425 430	
cag ggc aag agc atc gtc aac tcg atc agc ctc aag gag ggc gag gag	1344
Gln Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu	
435 440 445	
ctg ttc aag gag cgc gct cgc aag atc atg caa tac ggc gcg gcg gcg	1392
Leu Phe Lys Glu Arg Ala Arg Lys Ile Met Gln Tyr Gly Ala Ala Ala	
450 455 460	
gtg gtc atg gcc ttc gac gag cag ggc cag gcc gac agc ctg cac cgc	1440
Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Ser Leu His Arg	
465 470 475 480	
cgc atc gag att tgc agc cgc gcc tac aaa att ctc acc gaa gag gtg	1488
Arg Ile Glu Ile Cys Ser Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val	
485 490 495	

226

ggc ttc ccg ccg gag gac atc atc ttt gac ccg aac gtg ctg acc gtg	1536
Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val	
500 505 510	
gcc acc ggc atc gac gag cac aac aac tac gcg ctc gac ttc atc gaa	1584
Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile Glu	
515 520 525	
agc gtg ccg tgg atc aag cag aac ctg ccg cac gcg aag gtc tcc ggc	1632
Ser Val Arg Trp Ile Lys Gln Asn Leu Pro His Ala Lys Val Ser Gly	
530 535 540	
ggc atc agc aac gtt tcg ttc tcc ttc cgc ggc aac gag ccg gtg cgc	1680
Gly Ile Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg	
545 550 555 560	
gag gcg atg cac acc gcg ttc ctc tac cac gcc atc cac gcc ggt ctc	1728
Glu Ala Met His Thr Ala Phe Leu Tyr His Ala Ile His Ala Gly Leu	
565 570 575	
gac atg ggc atc gtc aac gcc gcc cag ctt ggc atc tac gaa gag atc	1776
Asp Met Gly Ile Val Asn Ala Ala Gln Leu Gly Ile Tyr Glu Glu Ile	
580 585 590	
gac ccg gag ctt ctt gtc tat gtc gag gac gtg ctg ctg aac cgc cgc	1824
Asp Pro Glu Leu Leu Val Tyr Val Glu Asp Val Leu Leu Asn Arg Arg	
595 600 605	
gac gac gcc acc gag cgg ctc gtg gcg ttc gct gaa acg atc cgc gac	1872
Asp Asp Ala Thr Glu Arg Leu Val Ala Phe Ala Glu Thr Ile Arg Asp	
610 615 620	
ggc ggc gaa aag gcc gag gcc aag aac gcc gaa tgg cgc aac gcc ccg	1920
Gly Gly Glu Lys Ala Glu Ala Lys Asn Ala Glu Trp Arg Asn Ala Pro	
625 630 635 640	
gtc gag gag cgg ctg aaa cac gcg ctc gtc aag ggc atc gtt gac tac	1968
Val Glu Glu Arg Leu Lys His Ala Leu Val Lys Gly Ile Val Asp Tyr	
645 650 655	
atc gac gag gac acc gaa gag gcc cgc cag ctc tac ccg agt ccg ctg	2016
Ile Asp Glu Asp Thr Glu Glu Ala Arg Gln Leu Tyr Pro Ser Pro Leu	
660 665 670	
gag gtg atc gag ggg ccg ctc atg aac ggc atg aac cac gtc ggc gac	2064
Glu Val Ile Glu Gly Pro Leu Met Asn Gly Met Asn His Val Gly Asp	
675 680 685	
ctc ttc gcc gaa ggc aag atg ttc ctg cca cag gtg gtc aaa agc gcc	2112
Leu Phe Ala Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala	
690 695 700	
cgc gtc atg aag cgc tcg gta gct gcg ctg att ccc tat atc gag gag	2160
Arg Val Met Lys Arg Ser Val Ala Ala Leu Ile Pro Tyr Ile Glu Glu	
705 710 715 720	
gag aag tcg aaa aac tgc gac acg agc gcc aaa gcc aag gtg ctg ctc	2208
Glu Lys Ser Lys Asn Cys Asp Thr Ser Ala Lys Ala Lys Val Leu Leu	
725 730 735	
gcc acg gtg aag ggc gac gtg cac gac atc ggc aag aac atc gtg tcg	2256



227

Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser	
740 745 750	
gtg gtg ctt gcc tgc aac aac ttc gac gtg atc gac atc ggc gtc atg	2304
Val Val Leu Ala Cys Asn Asn Phe Asp Val Ile Asp Ile Gly Val Met	
755 760 765	
atg cca tgc gac aag att ctc gaa gcg ctg gca gaa cac aag ccc gac	2352
Met Pro Cys Asp Lys Ile Leu Glu Ala Leu Ala Glu His Lys Pro Asp	
770 775 780	
gtg ctc ggc ctc tcc ggc ctc atc acc ccg tcg ctc gaa gag atg gcg	2400
Val Leu Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala	
785 790 795 800	
cac gtg gcc aaa gag atg gag cgg ctc ggc atg aac att ccg ctc atc	2448
His Val Ala Lys Glu Met Glu Arg Leu Gly Met Asn Ile Pro Leu Ile	
805 810 815	
atc ggc ggc gcg acc acc tcg aag gtg cac acg gcg gtg aaa ctc gcg	2496
Ile Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Leu Ala	
820 825 830	
ccc tgc tac ccc agc ggc gcg gta gta cac gtg ctc gac gcc tcg cgc	2544
Pro Cys Tyr Pro Ser Gly Ala Val Val His Val Leu Asp Ala Ser Arg	
835 840 845	
agc gtg ccg gtg gtc agc aac ctc tgc aac ccc gcc cag cgc gac agc	2592
Ser Val Pro Val Val Ser Asn Leu Cys Asn Pro Ala Gln Arg Asp Ser	
850 855 860	
tat atc gcg gcg ctg aag gat gag cag gag gcg atg cgc aag agc cac	2640
Tyr Ile Ala Ala Leu Lys Asp Glu Gln Glu Ala Met Arg Lys Ser His	
865 870 875 880	
gcc gag cgc atg gcg gca aaa aag tac gtc tcg ctc gac gcc gcc cgc	2688
Ala Glu Arg Met Ala Ala Lys Lys Tyr Val Ser Leu Asp Ala Ala Arg	
885 890 895	
gac aac cgc ctc acc att gac tgg gag gcc gaa acc atc gac aag ccc	2736
Asp Asn Arg Leu Thr Ile Asp Trp Glu Ala Glu Thr Ile Asp Lys Pro	
900 905 910	
gcc cag act ggc gtc acc gtg ctg gag gat gtc acc gtc ggc gcg ctc	2784
Ala Gln Thr Gly Val Thr Val Leu Glu Asp Val Thr Val Gly Ala Leu	
915 920 925	
cgc ccg tat atc gac tgg gca mcc ttc ttc tgg agc tgg gag ctg cac	2832
Arg Pro Tyr Ile Asp Trp Ala Xaa Phe Phe Trp Ser Trp Glu Leu His	
930 935 940	
ggc gtc tat ccg cag att ctg gag gat gaa aag gtc ggc gag gag gca	2880
Gly Val Tyr Pro Gln Ile Leu Glu Asp Glu Lys Val Gly Glu Glu Ala	
945 950 955 960	
acc aaa ctc ttc aac gac gcc acc gct ctg ctc gac cgg atc gac agc	2928
Thr Lys Leu Phe Asn Asp Ala Thr Ala Leu Leu Asp Arg Ile Asp Ser	
965 970 975	
gaa aag ctg ctc ggc atc aaa ggc gtg gcg ggc atc ttc ccg gcc aac	2976
Glu Lys Leu Leu Gly Ile Lys Gly Val Ala Gly Ile Phe Pro Ala Asn	
980 985 990	

228

agc atc ggc gac gac atc ttc gtc tat gcg gat gac gag cgc tcg ata 3024  
 Ser Ile Gly Asp Asp Ile Phe Val Tyr Ala Asp Asp Glu Arg Ser Ile  
 995 1000 1005

atc cgc acc gtg ctg cac acc ctg cgc cag caa ggc gaa aag cac ggc 3072  
 Ile Arg Thr Val Leu His Thr Leu Arg Gln Gln Gly Glu Lys His Gly  
 1010 1015 1020

gaa gcg aac ctc gcg ctg gcg gac ttc gtg gcc ccg cgc gaa agc ggc 3120  
 Glu Ala Asn Leu Ala Leu Ala Asp Phe Val Ala Pro Arg Glu Ser Gly  
 1025 1030 1035 1040

gtc aac gac tgg atc ggc tgc ttc acc gta acc gcc gga ctc ggc atc 3168  
 Val Asn Asp Trp Ile Gly Cys Phe Thr Val Thr Ala Gly Leu Gly Ile  
 1045 1050 1055

cag aat ttg ctc gac gag ttc aca gca gag aac gac gac tac cac cgc 3216  
 Gln Asn Leu Leu Asp Glu Phe Thr Ala Glu Asn Asp Asp Tyr His Arg  
 1060 1065 1070

atc atg aca cag gcg ctc gcc gac cga ctg gcc gaa gcg ttc gca gag 3264  
 Ile Met Thr Gln Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu  
 1075 1080 1085

atg ctg cac gaa aag gtg cgc cgc gaa ctc tgg ggc tac gcg ccc ggc 3312  
 Met Leu His Glu Lys Val Arg Arg Glu Leu Trp Gly Tyr Ala Pro Gly  
 1090 1095 1100

gaa atc ctc ggc aac gaa gag ctg atc gcc gaa aag tac cga ggc atc 3360  
 Glu Ile Leu Gly Asn Glu Glu Leu Ile Ala Glu Lys Tyr Arg Gly Ile  
 1105 1110 1115 1120

cgc ccc gcc ccc ggc tac ccc gcc tgc ccg gat cac acc gaa aag gca 3408  
 Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys Ala  
 1125 1130 1135

atc atc ttc gac ctg ctc aac gct gaa gcg gcc acc ggc gtc acg ctg 3456  
 Ile Ile Phe Asp Leu Leu Asn Ala Glu Ala Ala Thr Gly Val Thr Leu  
 1140 1145 1150

acg gaa act ttc gcg atg aac ccc gca gcc tca gtc tgc ggc ctc tac 3504  
 Thr Glu Thr Phe Ala Met Asn Pro Ala Ala Ser Val Cys Gly Leu Tyr  
 1155 1160 1165

ttc gcc aac ccg gcc tcg aaa tac ttc gta ctc ggc aag att ggt aag 3552  
 Phe Ala Asn Pro Ala Ser Lys Tyr Phe Val Leu Gly Lys Ile Gly Lys  
 1170 1175 1180

gat cag gtc gaa gac tac gcc aac cgc aaa ggg ctg gaa gta gca gaa 3600  
 Asp Gln Val Glu Asp Tyr Ala Asn Arg Lys Gly Leu Glu Val Ala Glu  
 1185 1190 1195 1200

gcc gag aag tgg ctc gcg ccc tcg ctg aac tac gat cca gcg 3642  
 Ala Glu Lys Trp Leu Ala Pro Ser Leu Asn Tyr Asp Pro Ala  
 1205 1210

taa 3645

&lt;210&gt; 52

&lt;211&gt; 1214

229

&lt;212&gt; PRT

&lt;213&gt; Chlorobium tepidum

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 936 .. 936

&lt;223&gt; All occurrences of Xaa indicate any amino acid

&lt;400&gt; 52

Val Leu Asp Gly Ala Met Gly Thr Met Ile Gln Arg His Gly Leu Asp  
 1 5 10 15

Glu Gln Asp Tyr Arg Gly Glu Arg Phe Ala Ser His Asp His Pro Leu  
 20 25 30

Lys Gly Asn Asn Asp Leu Leu Val Ile Thr Arg Pro Asp Ile Ile Arg  
 35 40 45

Ser Ile His Cys Asp Phe Leu Asp Ala Gly Ala Asp Ile Ile Glu Thr  
 50 55 60

Cys Thr Phe Asn Ala Asn Pro Ile Ser Gln Ser Asp Tyr Gln Leu Gln  
 65 70 75 80

Asp Leu Thr Arg Glu Leu Asn Val Ala Ala Ala Lys Ile Ala Arg Ser  
 85 90 95

Ala Ala Asp Glu Phe Thr Ala Lys Thr Pro Asp Lys Pro Arg Phe Val  
 100 105 110

Ala Gly Ser Ile Gly Pro Thr Asn Lys Thr Leu Ser Leu Ser Pro Asp  
 115 120 125

Val Asn Asn Pro Gly Phe Arg Ala Val Thr Phe Gln Glu Met Val Asp  
 130 135 140

Asn Tyr Thr Ala Gln Leu Glu Gly Leu His Glu Gly Gly Val Asp Leu  
 145 150 155 160

Leu Leu Val Glu Thr Val Phe Asp Thr Leu Asn Cys Lys Ala Ala Leu  
 165 170 175

Tyr Ala Ile Glu Glu Tyr Ala Val Lys Thr Gly Trp Gln Val Pro Val  
 180 185 190

Met Val Ser Gly Thr Val Val Asp Ala Ser Gly Arg Thr Leu Ser Gly  
 195 200 205

Gln Thr Thr Glu Ala Phe Trp Ile Ser Ile Ser His Met Pro Ser Leu  
 210 215 220

Leu Ser Val Gly Leu Asn Cys Ala Leu Gly Ser Lys Gln Met Arg Pro  
 225 230 235 240

Phe Ile Glu Ala Leu Ser Asn Ile Ala Glu Ser Tyr Val Ser Val Tyr  
 245 250 255

Pro Asn Ala Gly Leu Pro Asn Glu Phe Gly Glu Tyr Asp Asp Ser Pro  
 260 265 270

Glu Tyr Met Ala Ala Gln Ile Ala Gly Phe Ala Glu Ser Gly Phe Val  
 275 280 285

230

Asn Ile Val Gly Gly Cys Cys Gly Thr Thr Pro Thr His Ile Arg Ala  
 290 295 300  
 Ile Ala Glu Ala Val Lys Thr Leu Pro Pro Arg Lys Arg Pro Ala Asn  
 305 310 315 320  
 Lys His Val Leu Arg Leu Ser Gly Leu Glu Pro Leu Val Val Asp Glu  
 325 330 335  
 Thr Thr Gly Phe Ile Asn Val Gly Glu Arg Thr Asn Val Thr Gly Ser  
 340 345 350  
 Arg Lys Phe Ala Arg Leu Ile Lys Glu Ala Asn Tyr Asp Glu Ala Leu  
 355 360 365  
 Ser Ile Ala Arg Gln Gln Val Glu Asn Gly Ala Gln Val Ile Asp Val  
 370 375 380  
 Asn Leu Asp Glu Gly Met Leu Asp Ser Glu Lys Val Ile Val Glu Phe  
 385 390 395 400  
 Leu Asn Leu Ile Ala Ser Glu Pro Glu Ile Ala Lys Val Pro Val Met  
 405 410 415  
 Ile Asp Ser Ser Lys Trp Ser Val Ile Glu Asn Gly Leu Arg Cys Thr  
 420 425 430  
 Gln Gly Lys Ser Ile Val Asn Ser Ile Ser Leu Lys Glu Gly Glu Glu  
 435 440 445  
 Leu Phe Lys Glu Arg Ala Arg Lys Ile Met Gln Tyr Gly Ala Ala Ala  
 450 455 460  
 Val Val Met Ala Phe Asp Glu Gln Gly Gln Ala Asp Ser Leu His Arg  
 465 470 475 480  
 Arg Ile Glu Ile Cys Ser Arg Ala Tyr Lys Ile Leu Thr Glu Glu Val  
 485 490 495  
 Gly Phe Pro Pro Glu Asp Ile Ile Phe Asp Pro Asn Val Leu Thr Val  
 500 505 510  
 Ala Thr Gly Ile Asp Glu His Asn Asn Tyr Ala Leu Asp Phe Ile Glu  
 515 520 525  
 Ser Val Arg Trp Ile Lys Gln Asn Leu Pro His Ala Lys Val Ser Gly  
 530 535 540  
 Gly Ile Ser Asn Val Ser Phe Ser Phe Arg Gly Asn Glu Pro Val Arg  
 545 550 555 560  
 Glu Ala Met His Thr Ala Phe Leu Tyr His Ala Ile His Ala Gly Leu  
 565 570 575  
 Asp Met Gly Ile Val Asn Ala Ala Gln Leu Gly Ile Tyr Glu Glu Ile  
 580 585 590  
 Asp Pro Glu Leu Leu Val Tyr Val Glu Asp Val Leu Leu Asn Arg Arg  
 595 600 605  
 Asp Asp Ala Thr Glu Arg Leu Val Ala Phe Ala Glu Thr Ile Arg Asp

231

610	615	620
Gly Gly Glu Lys Ala Glu Ala Lys Asn Ala Glu Trp Arg Asn Ala Pro		
625	630	635 640
Val Glu Glu Arg Leu Lys His Ala Leu Val Lys Gly Ile Val Asp Tyr		
	645	650 655
Ile Asp Glu Asp Thr Glu Glu Ala Arg Gln Leu Tyr Pro Ser Pro Leu		
	660	665 670
Glu Val Ile Glu Gly Pro Leu Met Asn Gly Met Asn His Val Gly Asp		
	675	680 685
Leu Phe Ala Glu Gly Lys Met Phe Leu Pro Gln Val Val Lys Ser Ala		
	690	695 700
Arg Val Met Lys Arg Ser Val Ala Ala Leu Ile Pro Tyr Ile Glu Glu		
	705	710 715 720
Glu Lys Ser Lys Asn Cys Asp Thr Ser Ala Lys Ala Lys Val Leu Leu		
	725	730 735
Ala Thr Val Lys Gly Asp Val His Asp Ile Gly Lys Asn Ile Val Ser		
	740	745 750
Val Val Leu Ala Cys Asn Asn Phe Asp Val Ile Asp Ile Gly Val Met		
	755	760 765
Met Pro Cys Asp Lys Ile Leu Glu Ala Leu Ala Glu His Lys Pro Asp		
	770	775 780
Val Leu Gly Leu Ser Gly Leu Ile Thr Pro Ser Leu Glu Glu Met Ala		
	785	790 795 800
His Val Ala Lys Glu Met Glu Arg Leu Gly Met Asn Ile Pro Leu Ile		
	805	810 815
Ile Gly Gly Ala Thr Thr Ser Lys Val His Thr Ala Val Lys Leu Ala		
	820	825 830
Pro Cys Tyr Pro Ser Gly Ala Val His Val Leu Asp Ala Ser Arg		
	835	840 845
Ser Val Pro Val Val Ser Asn Leu Cys Asn Pro Ala Gln Arg Asp Ser		
	850	855 860
Tyr Ile Ala Ala Leu Lys Asp Glu Gln Glu Ala Met Arg Lys Ser His		
	865	870 875 880
Ala Glu Arg Met Ala Ala Lys Lys Tyr Val Ser Leu Asp Ala Ala Arg		
	885	890 895
Asp Asn Arg Leu Thr Ile Asp Trp Glu Ala Glu Thr Ile Asp Lys Pro		
	900	905 910
Ala Gln Thr Gly Val Thr Val Leu Glu Asp Val Thr Val Gly Ala Leu		
	915	920 925
Arg Pro Tyr Ile Asp Trp Ala Xaa Phe Phe Trp Ser Trp Glu Leu His		
	930	935 940

232

Gly Val Tyr Pro Gln Ile Leu Glu Asp Glu Lys Val Gly Glu Glu Ala  
 945 950 955 960  
 Thr Lys Leu Phe Asn Asp Ala Thr Ala Leu Leu Asp Arg Ile Asp Ser  
 965 970 975  
 Glu Lys Leu Leu Gly Ile Lys Gly Val Ala Gly Ile Phe Pro Ala Asn  
 980 985 990  
 Ser Ile Gly Asp Asp Ile Phe Val Tyr Ala Asp Asp Glu Arg Ser Ile  
 995 1000 1005  
 Ile Arg Thr Val Leu His Thr Leu Arg Gln Gln Gly Glu Lys His Gly  
 1010 1015 1020  
 Glu Ala Asn Leu Ala Leu Ala Asp Phe Val Ala Pro Arg Glu Ser Gly  
 1025 1030 1035 1040  
 Val Asn Asp Trp Ile Gly Cys Phe Thr Val Thr Ala Gly Leu Gly Ile  
 1045 1050 1055  
 Gln Asn Leu Leu Asp Glu Phe Thr Ala Glu Asn Asp Asp Tyr His Arg  
 1060 1065 1070  
 Ile Met Thr Gln Ala Leu Ala Asp Arg Leu Ala Glu Ala Phe Ala Glu  
 1075 1080 1085  
 Met Leu His Glu Lys Val Arg Arg Glu Leu Trp Gly Tyr Ala Pro Gly  
 1090 1095 1100  
 Glu Ile Leu Gly Asn Glu Glu Leu Ile Ala Glu Lys Tyr Arg Gly Ile  
 1105 1110 1115 1120  
 Arg Pro Ala Pro Gly Tyr Pro Ala Cys Pro Asp His Thr Glu Lys Ala  
 1125 1130 1135  
 Ile Ile Phe Asp Leu Leu Asn Ala Glu Ala Ala Thr Gly Val Thr Leu  
 1140 1145 1150  
 Thr Glu Thr Phe Ala Met Asn Pro Ala Ala Ser Val Cys Gly Leu Tyr  
 1155 1160 1165  
 Phe Ala Asn Pro Ala Ser Lys Tyr Phe Val Leu Gly Lys Ile Gly Lys  
 1170 1175 1180  
 Asp Gln Val Glu Asp Tyr Ala Asn Arg Lys Gly Leu Glu Val Ala Glu  
 1185 1190 1195 1200  
 Ala Glu Lys Trp Leu Ala Pro Ser Leu Asn Tyr Asp Pro Ala  
 1205 1210

&lt;210&gt; 53

&lt;211&gt; 52

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; Beschreibung der künstlichen Sequenz:PCR primer

&lt;400&gt; 53

ccgggatcc gctagcggcg cgccggccgg cccggtgtga aataccgcac ag

52

<210> 54  
<211> 53  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 54  
tctagactcg agcgggccgcg gccggccttt aaattgaaga cgaaagggcc tcg 53

<210> 55  
<211> 47  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 55  
gagatctaga cccgggggatc cgctagcggg ctgctaaagg aagcgga 47

<210> 56  
<211> 38  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 56  
gagagggcgcg ccgctagcgt gggcgaagaa ctccagca 38

<210> 57  
<211> 34  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 57  
gagagggcgcg ccgcgcaaag tcccgttcg tgaa 34

<210> 58  
<211> 34  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 58  
gagagggcgcg ccgctcaagt cggtaagcc acgc 34

<210> 59  
<211> 140  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 59  
tcgaatttaa atctcgagag gcctgacgtc gggcccggtta ccacgcgtca tatgactagt 60  
tcggacctag ggatatacgtc gacatcgatg ctcttctgcg ttaattaaca attgggatcc 120  
tctagaccgc ggatttaaatt 140

<210> 60  
<211> 140  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 60  
gatcatttaa atccccgggtc tagaggatcc caattgttaa ttaacgcaga agagcatcga 60  
tgtcgacgat atccctaggt ccgaactagt catatgacgc gtggtaccgg gcccgacgtc 120  
aggcctctcg agatttaaatt 140

<210> 61  
<211> 33  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 61  
gagagcggcc gccgatacctt ttttaacccat cac 33

<210> 62  
<211> 32  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 62  
aggagcggcc gccatcgga ttttcttttg cg 32

<210> 63  
<211> 5091  
<212> DNA  
<213> Künstliche Sequenz

<220>  
<223> Beschreibung der künstlichen Sequenz:Plasmid

<400> 63  
gccgcgactg ccttcgcgaa gccttgcccc gcggaaattt cctccaccga gttcgtgcac 60



235

acccctatgc caagcttctt tcaccctaaa ttcgagagat tggattctta ccgtggaat 120  
 tcttcgcaaa aatcgctccc tgatcgccct tgcgacgttg gcgtcgggtgc cgctgggtgc 180  
 gcttggcttg accgacttga tcagcggccg ctcgatttaa atctcgagag gcctgacgtc 240  
 gggcccggtt ccacgcgtca tatgactagt tcggacctag ggatatcgtc gacatcgatg 300  
 ctcttctgcg ttaattaaca attgggatcc tctagaccgg ggatttaa at cgctagcggg 360  
 ctgctaaagg aagcggaa ca tagaaagc cagtccgcag aaacgggtgct gaccccggt 420  
 gaatgtcagc tactgggcta tctggacaag ggaaaacgca agcgcgaaaga gaaagcaggt 480  
 agcttgacgt gggcttacat ggcgatagct agactggggc gttttatgga cagcaagcga 540  
 accggaattg ccagctgggg cgccctctgg taagggtggg aagccctgca aagtaaaactg 600  
 gatggcttct ttgcccga ggatctgat gcgcagggga tcaagatctg atcaagagac 660  
 aggatgagga tcgtttcgca tgattgaaca agatggattg cacgcaggtt ctccggccgc 720  
 ttgggtggag aggtattcg gctatgactg ggcacaacag acaatcggct gctctgatgc 780  
 cgcgtgttc cggctgtcag cgcagggggc cccggttctt tttgtcaaga ccgacctgtc 840  
 cggcgccctg aatgaactgc aggcagggc agcgcggcta tcgtgggtgg ccacgacggg 900  
 cgttccttgc gcagctgtgc tcgacgttgt cactgaagcg ggaagggaact ggctgctatt 960  
 gggcgaaagt ccggggcagg atctcctgtc atctcactt gctcctgccc agaaagtatc 1020  
 catcatggct gatgcaatgc ggcggctgca tacgcttgat ccggctacct gccatttca 1080  
 ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg atggaagccg gtcttgcga 1140  
 tcaggatgat ctggacgaag agcatcaggg gctcgcgcca gccgaactgt tcgccaggct 1200  
 caaggcgcgc atgcccgcag gcgaggatct cgtcgtgacc catggcgatg cctgcttgc 1260  
 gaatatcatg gtggaaaatg gccgcttttc tggattcatc gactgtggcc ggctgggtgt 1320  
 ggcggaccgc tatcaggaca tagcgttggc taccgctgat attgtgaag agcttggcgg 1380  
 cgaatgggct gaccgcttcc tcgtgtttaa cggtatcgcc gctcccgatt cgcagcgc 1440  
 cgcttctat cgcttcttg acgagttctt ctgagcggga ctctggggtt cgaatgacc 1500  
 gaccaagcga cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa 1560  
 aggttgggct tcggaatcgt tttccgggac gccggctgga tgatcctcca gcgcggggat 1620  
 ctcatgctgg agttcttcgc ccacgctagc ggcgcccggt ccggcccggt gtgaaatacc 1680  
 gcacagatgc gtaaggagaa aataccgcac caggcgctct tccgcttctt cgtcactga 1740  
 ctcgctgcgc tcggtcgttc ggctgcggcg agcggatatca gctcactcaa aggcggaat 1800  
 acggttatcc acagaatcag gggataacgc aggaagaac atgtgagcaa aaggccagca 1860  
 aaaggccagg aaccgtaaaa aggcgcggtt gctggcggtt ttccataggc tcgcccccc 1920  
 tgacgagcat caaaaaatc gacgtcaag tcagaggtgg cgaaacccga caggactata 1980  
 aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgacctgcc 2040  
 gcttacccga tacctgtccg ccttctccc tcgggaagc gtggcgcttt ctcatagctc 2100  
 acgctgtagg tatctcagtt cgggttaggt cgttcgctcc aagctgggct gtgtgcaga 2160  
 acccccgtt cagcccagac gctgcgcctt atccgtaac tatcgtcttg agtccaacce 2220  
 ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 2280  
 gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct aactagaag 2340  
 gacagtattt ggtatctgag ctctgctgaa gccagttacc ttcggaaaaa gagttggtag 2400  
 ctcttgatcc ggcaaacaaa ccaccgctgg tagcgttggg tttttgttt gcaagcagca 2460  
 gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 2520  
 cgctcagtg aacgaaaact cacgttaagg gattttgttc atgagattat caaaaaggat 2580  
 cttcacctag atccttttaa aggccggcg cggccgcgca aagtcgcgt tcgtgaaaat 2640  
 tttcgtgcg cgtgattttc gcgcaaaaac tttaacgaac gttcgttata atggtgtcat 2700  
 gaccttcacg acgaagtact aaaattggcc cgaatcatca gctatggatc tctctgatgt 2760  
 cgcgtggag tccgacgcgc tcgatgctgc cgtcgattta aaaacgggtga tcggattttt 2820  
 ccgagctctc gatacgacgg acgcgccagc atcacgagac tgggccaagt ccgcgagcga 2880  
 cctagaaact ctcgtggcg atcttgagga gctggctgac gagctgcgtg ctccggccagc 2940  
 gccaggagga cgcacagtag tggaggatgc aatcagttgc gcctactgcg gtggcctgat 3000  
 tctccccgg cctgaccgc gaggacggcg cgcaaaatat tgctcagatg cgtgtcgtgc 3060  
 cgcagccagc cgcgagcgcg ccaacaaaac ccacgccgag gagctggagg cggttaggtc 3120  
 gcaaatggcg ctggaagtgc gtccccgag cgaaattttg gccatggtcg tcacagagct 3180  
 ggaagcggca gcgagaatta tcgcatcgt ggcgggtgcc gcaggcatga caaacatcgt 3240  
 aaatgccgcg tttcgtgtgc cgtggccgcc caggacgtgt cagcgccgc accacctgca 3300  
 ccgaatcggc agcagcgtcg cgcgtcgaaa aagcgcacag gcggcaagaa gcgataagct 3360  
 gcacgaatac ctgaaaaatg ttgaacgccc cgtgagcgg aactcacagg gcgtcggcta 3420  
 acccccagtc caaacctggg agaaagcgt caaaaatgac tctagcggat tcacgagaca 3480  
 ttgacacacc ggcctggaaa ttttccgctg atctgttca caccatccc gagctcgcgc 3540  
 tgcgatcagc tggctggagc agcgaagacc gccgcgaatt cctcgtcac ctggcgagag 3600  
 aaaatttcca gggcagcaag acccgcgact tcgcccagcg ttggatcaaa gaccgggaca 3660  
 cggagaaaca cagccgaagt tataccgagt tggttcaaaa tcgcttgccc ggtgccagta 3720  
 tgttgcctc agcacgcgc agcacgcagc cgtgcttgc ctggacattg atgtgccag 3780

236

```

ccaccaggcc ggcgggaaaa tcgagcacgt aaaccccgag gtctacgcga ttttggagcg 3840
ctggggcacgc ctggaaaaag cgcagccttg gatcggcgtg aatccactga gcgggaaatg 3900
ccagctcatc tggctcattg atccgggtga tgccgcagca ggcatgagca gcccgaatat 3960
gcgcctgctg gctgcaacga ccgaggaaat gacccgcgtt ttcggcgtg accaggcttt 4020
ttcacatagg ctgagccgtg gccactgcac tctccgacga tcccagccgt accgctggca 4080
tgcccagcac aatcgcggtg atcgccatga gcaggagttt tctagcggac gggcacgtat 4200
aggcacagaa aaacctaata aacgctatga gcaggagttt tctagcggac gggcacgtat 4200
cgaagcggca agaaaaagcca ctgcggaagc aaaagcactt gccacgcttg aagcaagcct 4260
gccgagcgcc gctgaagcgt ctggagagct gatcgacggc gtccgtgtcc tctggactgc 4320
tccaggcgct gccgccgtg atgagacggc ttttcgccac gctttgactg tgggatacca 4380
gttaaaagcg gctggtgagc gcctaaaaga caccaagggt catcgagcct acgagcgtgc 4440
ctacaccgtc gctcaggcgg tcggaggagg ccgtgagcct gatctgccgc cggactgtga 4500
ccgccagacg gattggccgc gacgtgtgcg cggctacgtc gctaaaggcc agccagtcgt 4560
ccctgctcgt cagacagaga cgcagagcca gccgaggcga aaagctcttg ccactatggg 4620
aagacgtggc ggtaaaaagg ccgcagaacg ctggaaagac ccaaacagtg agtacgccc 4680
agcacagcga gaaaaactag ctaagtccag tcaacgacaa gctaggaaag ctaaaaggaaa 4740
tcgcttgacc attgcagggt ggtttatgac tgttgaggga gagactggct cgtggccgac 4800
aatcaatgaa gctatgtctg aatttagcgt gtcacgtcag accgtgaata gagacttaa 4860
ggctcgcggg cattgaactt ccacgaggac gccgaaagct tcccagtaaa tgtgccatct 4920
cgtaggcaga aaacggttcc ccgtaggggt ctctctcttg gcctccttcc taggtcgggc 4980
tgattgctct taagctcttc taggggggct cacaccatag gcagataacg tccccaccg 5040
gctcgcctcg taagcgcaca aggactgctc ccaaagatct tcaaagccac t 5091

```

&lt;210&gt; 64

&lt;211&gt; 4323

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; Beschreibung der künstlichen Sequenz: Plasmid

&lt;400&gt; 64

```

tctctcagcg tatggttgct gcctgagctg tagttgcctt catcgatgaa ctgctgtaca 60
ttttgatacg tttttccgtc accgtcaaaag attgatattat aatccctctac accgttgatg 120
ttcaaagagc tgcctgatgc tgatacgtta acttggtgcag ttgtcagtggt ttgtttgccg 180
taatgtttac cggagaaatc agtgtagaat aaacggattt ttccgtcaga tgtaaagtgt 240
gctgaacctg accattcttg tgtttggtct tttaggatag aatcatttgc atcgaatttg 300
tcgctgtctt taaagacgcg gccagcgttt ttccagctgt caatagaagt ttccgccgact 360
ttttgataga acatgtaaat cgatgtgtca tccgcatttt taggatctcc ggctaatagca 420
aagacgatgt ggtagccgtg atagtttgcg acagtgcctt cagcgttttg taatggccag 480
ctgtcccaaa cgtccaggcc ttttcgagaa gagatatttt taattgtgga cgaatcaaat 540
tcagaaactt gatatttttc atttttttgc tgttcaggga ttgacagcat atcatggcgt 600
gtaatatggg aaatgccgta tgtttcctta tatggccttt gggtcgtttc ttccgcaaac 660
gcttgagttg cgctcctgc cagcagtgcg gtagtaaagg ttaatactgt tgcttgtttt 720
gcaaaacttt tgatgttcat cgttcatgtc tcttttttta tgtactgtgt tagcggctctg 780
cttcttccag cctcctgtt tgaagatggc aagttagtta cgcacaataa aaaaagacct 840
aaaatatgta aggggtgacg ccaaagtata cactttgcc tttacacatt ttaggtcttg 900
cctgctttat cagtaacaaa ccgcgcgat ttacttttcg acctcattct attagactct 960
cgtttggatt gcaactggtc tattttcttc tttgttttga tagaaaaatc taaaaggatt 1020
tgcagactac gggcctaaag aactaaaaaa tctatctggt tcttttcatt ctctgtattt 1080
tttatagttt ctggtgcatg ggcataaagt tgctttttta atcacaattc agaaaatata 1140
ataatatctc atttactaa ataatagtga acggcaggtat tatgtgatgg gttaaaaagg 1200
atcgccggcc gctcgattta aatctcgaga ggcctgacgt cgggcccggg accacgcgtc 1260
atatgactag ttcggacctt gggatatcgt cgacatcgat gctcttctgc gttaattaac 1320
aattgggata ctctagacc cggattttaa tcgctagcgg gctgctaaag gaagcggaac 1380
acgtagaaag ccagtcgca gaaacgggtg tgaccccgga tgaatgtcag ctactgggct 1440
atctggacaa gggaaaacgc aagcgcgaa agaaagcagg tagcttgacg tgggcttaca 1500
tggcgatagc tagactgggc ggttttatgg acagcaagcg aaccggaatt gccagctggg 1560
gcgcctctg gtaaggttg gaagccctgc aaagtaaaact ggatggcttt cttgccgcca 1620
aggatctgat ggcgcagggg atcaagatct gatcaagaga caggatgagg atcgtttcgc 1680

```

## 237

atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggtatctc 1740  
 ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 1800  
 gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 1860  
 caggacgagg cagcgcgggt atcgtggctg gccacgcagg gcgttccttg cgcagctgtg 1920  
 ctgcagcttg tcaactgaagc gggaaggac tggtctgat tgggcgaagt gccggggcag 1980  
 gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 2040  
 cggcggtgc atacgcttga tccggctacc tgccattcgc accaccaagc gaaacatcgc 2100  
 atcgagcag cactactcg gatggaagcc ggtcttgctg atcaggatga tctggacgaa 2160  
 gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgac 2220  
 ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 2280  
 ggccgctttt ctggattcat cgaactgtggc cggctgggtg tggcggaccg ctatcaggac 2340  
 atagcgttgg ctaccctgta tattgctgaa gagcttgccg gcgaatggc tgaccgcttc 2400  
 ctgctgtttt acggtatcgc cgtcccgat tcgcagcgca tcgccttcta tcgccttctt 2460  
 gagcagttct tctgagcggg actctggggg tcgaaatgac cgaccaagcg acgccaacc 2520  
 tgccatcacg agatttcgat tccaccgcgc ccttctatga aagggtgggc ttcggaatcg 2580  
 ttttccggga cgcggctg atgatctctc agcgcgggga tctcatgctg gatttcttcg 2640  
 cccacgctag cggcgccgcg gccggcccg tgtgaaatac cgcacagatg cgtaggaga 2700  
 aaataccgca tcaggcgctc tccgcttcc tcgctcactg actcgtgcg ctccgctcgtt 2760  
 cggctgcggc gagcggatc agctcactca aaggcggtta tacggttatc cacagaatca 2820  
 ggggataaac caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 2880  
 aaggcgcgt tgctggcgtt ttccatagg ctcgccccc ctgacgagca tcacaaaaat 2940  
 cgacgctcaa gtcagaggtg gcgaaaccgc acaggactat aaagatacca ggcgtttccc 3000  
 cctggaagct cctcgtgctg ctctcctgtt ccgacctgc cgtttaccgg atacctgtcc 3060  
 gcctttctcc ctccgggaag cgtggcgctt tctcatagct cagctgttag gtatctcagt 3120  
 tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac 3180  
 cgctgcgct tatccggtta ctatcgtctt gactccaacc cggtaagaca cgacttatcg 3240  
 ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cgtgtctaca 3300  
 gagttcttga agtgggtggc taactacggc tacactagaa ggacagtatt tggatctcgc 3360  
 gctctgctga agccagttac ctccgaaaa agagtttgta gctcttgatc cggcaaaaaa 3420  
 accaccgctg gtacgggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 3480  
 ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg gaacgaaaaa 3540  
 tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 3600  
 aaggcggcc cggcgcccca tcggcatctt ctttgcgtt tttatttgtt aactgttaat 3660  
 tgccttgtt caaggatgct gtctttgaca acagatgttt tcttgccctt gatgttcagc 3720  
 aggaagctcg gcgcaaacgt tgattgtttg tctgcgtaga atcctctgtt tgtcatatag 3780  
 cttgtaatca cgacattgtt tccttgcgt tgaggtagag cgaagtgtga gtaagtaaag 3840  
 gttacatcgt taggatcaag atccattttt aacacaaggc cagttttgtt cagcggcttg 3900  
 tatgggccag ttaaagaatt agaaacataa ccaagcatgt aaatatcgtt agacgtaatg 3960  
 ccgtcaatcg tcatttttga tccgcgggag tcagtgaaca ggtaccattt gccgttcatt 4020  
 ttaaagacgt tcgcgcgttc aatttcactt gttactgtgt tagatgcaat cagcggtttc 4080  
 atcacttttt tcagtgtgta atcatcgtt agctcaatca taccgagagc gccgtttgct 4140  
 aactcagccg tgcgtttttt atcgctttgc agaagttttt gactttcttg acggaagaat 4200  
 gatgtgcttt tgccatagta tgctttgtta aataaagatt cttcgccttg gtagccatct 4260  
 tcagttccag tgtttgcttc aaatactaag tatttggtggc ctttatcttc tacgtagtga 4320  
 gga 4323

&lt;210&gt; 65

&lt;211&gt; 35

&lt;212&gt; DNA

&lt;213&gt; PCR Primer

&lt;400&gt; 65

gagagagaga cgcgtcccag tggctgagac gcac

35

&lt;210&gt; 66

&lt;211&gt; 34

&lt;212&gt; DNA

&lt;213&gt; PCR Primer

&lt;400&gt; 66

ctctctctgt cgacgaattc aatcttacgg cctg

34

<210> 67  
<211> 38  
<212> DNA  
<213> PCR Primer  
<400> 67

cggcaccacc gacatcatct tcacctgcc tcgttcg

38

<210> 68  
<211> 38  
<212> DNA  
<213> PCR Primer  
<400> 68

cggaaacgagg gcaggtgaag atgatgtcgg tggtgccg

38

<210> 69  
<211> 31  
<212> DNA  
<213> PCR Primer  
<400> 69

gagactcgag ggaaggtgaa tcgaatttcg g

31

<210> 70  
<211> 38  
<212> DNA  
<213> PCR Primer  
<400> 70

gtcccgggga gaacgcacga ttctccaaaa ataatcg

38

<210> 71  
<211> 23  
<212> DNA  
<213> PCR Primer  
<400> 71

gaatcgtgcg ttctccccgg gac

23

<210> 72  
<211> 22  
<212> DNA  
<213> PCR Primer  
<400> 72

gtagttgacc gagttgatca cc

22

<210> 73  
<211> 18  
<212> DNA  
<213> PCR Primer  
<400> 73

ccggcctgga gaagctcg

18

<210> 74  
<211> 28  
<212> DNA  
<213> PCR Primer  
<400> 74

239

gagagatatc cctcagcggg cgttgaag

28

<210> 75  
 <211> 1266  
 <212> DNA  
 <213> LysC Mutante  
 <220>  
 <221> CDS  
 <222> (1) .. (1266)  
 <223>

<400> 75  
 gtg gcc ctg gtc gta cag aaa tat ggc ggt tcc tcg ctt gag agt gcg 48  
 Val Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala  
 1 5 10 15  
 gaa cgc att aga aac gtc gct gaa cgg atc gtt gcc acc aag aag gct 96  
 Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala  
 20 25 30  
 gga aat gat gtc gtg gtt gtc tgc tcc gca atg gga gac acc acg gat 144  
 Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp  
 35 40 45  
 gaa ctt cta gaa ctt gca gcg gca gtg aat ccc gtt ccg cca gct cgt 192  
 Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg  
 50 55 60  
 gaa atg gat atg ctc ctg act gct ggt gag cgt att tct aac gct ctc 240  
 Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu  
 65 70 75 80  
 gtc gcc atg gct att gag tcc ctt ggc gca gaa gcc caa tct ttc acg 288  
 Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr  
 85 90 95  
 ggc tct cag gct ggt gtg ctc acc acc gag cgc cac gga aac gca cgc 336  
 Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg  
 100 105 110  
 att gtt gat gtc act cca ggt cgt gtg cgt gaa gca ctc gat gag ggc 384  
 Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly  
 115 120 125  
 aag atc tgc att gtt gct ggt ttc cag ggt gtt aat aaa gaa acc cgc 432  
 Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg  
 130 135 140  
 gat gtc acc acg ttg ggt cgt ggt ggt tct gac acc act gca gtt gcg 480  
 Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala  
 145 150 155 160  
 ttg gca gct gct ttg aac gct gat gtg tgt gag att tac tcg gac gtt 528  
 Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val  
 165 170 175  
 gac ggt gtg tat acc gct gac ccg cgc atc gtt cct aat gca cag aag 576  
 Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys  
 180 185 190  
 ctg gaa aag ctc agc ttc gaa gaa atg ctg gaa ctt gct gct gtt ggc 624

240

Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly	
195 200 205	
tcc aag att ttg gtg ctg cgc agt gtt gaa tac gct cgt gca ttc aat	672
Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn	
210 215 220	
gtg cca ctt cgc gta cgc tcg tct tat agt aat gat ccc ggc act ttg	720
Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu	
225 230 235 240	
att gcc ggc tct atg gag gat att cct gtg gaa gaa gca gtc ctt acc	768
Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr	
245 250 255	
ggg gtc gca acc gac aag tcc gaa gcc aaa gta acc gtt ctg ggt att	816
Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile	
260 265 270	
tcc gat aag cca ggc gag gct gcg aag gtt ttc cgt gcg ttg gct gat	864
Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp	
275 280 285	
gca gaa atc aac att gac atg gtt ctg cag aac gtc tct tct gta gaa	912
Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu	
290 295 300	
gac ggc acc acc gac atc atc ttc acc tgc cct cgt tcc gac ggc cgc	960
Asp Gly Thr Thr Asp Ile Ile Phe Thr Cys Pro Arg Ser Asp Gly Arg	
305 310 315 320	
cgc gcg atg gag atc ttg aag aag ctt cag gtt cag ggc aac tgg acc	1008
Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr	
325 330 335	
aat gtg ctt tac gac gac cag gtc ggc aaa gtc tcc ctc gtg ggt gct	1056
Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala	
340 345 350	
ggc atg aag tct cac cca ggt gtt acc gca gag ttc atg gaa gct ctg	1104
Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu	
355 360 365	
cgc gat gtc aac gtg aac atc gaa ttg att tcc acc tct gag att cgt	1152
Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg	
370 375 380	
att tcc gtg ctg atc cgt gaa gat gat ctg gat gct gct gca cgt gca	1200
Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala	
385 390 395 400	
ttg cat gag cag ttc cag ctg ggc ggc gaa gac gaa gcc gtc gtt tat	1248
Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr	
405 410 415	
gca ggc acc gga cgc taa	1266
Ala Gly Thr Gly Arg	
420	

<210> 76  
 <211> 421

241

<212> PRT  
 <213> LysC Mutante  
 <400> 76

Val Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala  
 1 5 10 15

Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala  
 20 25 30

Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp  
 35 40 45

Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg  
 50 55 60

Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu  
 65 70 75 80

Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr  
 85 90 95

Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg  
 100 105 110

Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly  
 115 120 125

Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg  
 130 135 140

Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala  
 145 150 155 160

Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val  
 165 170 175

Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys  
 180 185 190

Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly  
 195 200 205

Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn  
 210 215 220

Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu  
 225 230 235 240

242

Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr  
 245 250 255

Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile  
 260 265 270

Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp  
 275 280 285

Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu  
 290 295 300

Asp Gly Thr Thr Asp Ile Ile Phe Thr Cys Pro Arg Ser Asp Gly Arg  
 305 310 315 320

Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr  
 325 330 335

Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala  
 340 345 350

Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu  
 355 360 365

Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg  
 370 375 380

Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala  
 385 390 395 400

Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr  
 405 410 415

Ala Gly Thr Gly Arg  
 420

<210> 77

<211> 5860

<212> DNA

<213> Plasmid

<400> 77

cccgggtacca cgcgtcccag tggctgagac gcatccgcta aagccccagg aacctgtgc 60  
 agaaagaaaa cactcctctg gctaggtaga cacagtttat aaaggtagag ttgagcgggt 120  
 aactgtcagc acgtagatcg aaaggtgcac aaaggtggcc ctggtcgtac agaaatatgg 180  
 cggttcctcg cttgagagtg cggaacgcac tagaaacgtc gctgaacgga tcgttgccac 240



## 243

caagaaggct ggaaatgatg tcgtggttgt ctgctccgca atgggagaca ccacggatga 300  
acttctagaa cttgcagcgg cagtgaatcc cgttccgcca gctcgtgaaa tggatatgct 360  
cctgactgct ggtgagcgt tttctaacgc tctcgtcgcc atggctattg agtccccttg 420  
cgcagaagcc caatctttca cgggctctca ggctggtgtg ctcaccaccg agcgccacgg 480  
aaacgcacgc attgttgatg tcaactccagg tcgtgtgcgt gaagcactcg atgagggcaa 540  
gatctgcatt gttgctggtt tccaggggtg taataaagaa acccgcgatg tcaccacgtt 600  
gggtcgtggt ggttctgaca ccactgcagt tgcgttgga gctgctttga acgctgatgt 660  
gtgtgagatt tactcggacg ttgacgggtg gtataccgct gaccgcgca tcgttccctaa 720  
tgcacagaag ctggaaaagc tcagcttcga agaaatgctg gaacttgctg ctggttgctc 780  
caagattttg gtgctgcgca gtgttgaata cgctcgtgca ttcaatgtgc cacttcgcgt 840  
acgctcgtct tatagtaatg atccccgcac tttgattgcc ggctctatgg aggatattcc 900  
tgtggaagaa gcagtcctta ccggtgtcgc aaccgacaag tccgaagcca aagtaaccgt 960  
tctgggtatt tccgataagc caggcgaggc tgcgaagggt ttccgtgcgt tggctgatgc 1020  
agaaatcaac attgacatgg ttctgcagaa cgtctcttct gtagaagacg gcaccaccga 1080  
catcaccttc acctgccctc gttccgacgg ccgccgcg atggagatct tgaagaagct 1140  
tcaggttcag ggcaactgga ccaatgtgct ttacgacgac caggctcgga aagtctccct 1200  
cgtgggtgct ggcataaagt ctcaccacgg tgttaccgca gagttcatgg aagctctgcg 1260  
cgatgtcaac gtgaacatcg aattgatttc cacctctgag attcgtatct ccgtgctgat 1320  
ccgtgaagat gatctggatg ctgctgcacg tgcattgcat gagcagttcc agctgggcgg 1380  
cgaagacgaa gccgtcgttt atgcaggcac cggacgctaa agttttaag gagtagtttt 1440  
acaatgacca ccatacgagt tgttggtgca accggccagg tcggccagggt tatgcgacc 1500  
cttttggaag agcgcaattt cccagctgac actgttcgtt tctttgcttc cccacgttcc 1560  
gcaggccgta agattgaatt cgtcgacatc gatgctcttc tgcgttaatt aacaattggg 1620  
atcctctaga cccgggattt aaatcgctag cgggctgcta aaggaagcgg aacacgtaga 1680  
aagccagtcc gcagaaacgg tgetgacccc ggatgaatgt cagctactgg gctatctgga 1740  
caagggaana cgcaagcgca aagagaaagc aggtagcttg cagtgggctt acatggcgat 1800  
agctagactg ggcggtttta tggacagcaa gcgaaccgga attgccagct ggggcgcctt 1860  
ctggttaagg tgggaagccc tgcaaagtaa actggatggc tttcttgccg ccaaggatct 1920  
gatggcgag gggatcaaga tctgatcaag agacaggatg aggatcgctt cgcattgattg 1980  
aacaagatgg attgcacgca ggttctccgg ccgcttgggt ggagaggcta ttcggctatg 2040  
actgggcaca acagacaatc ggctgctctg atgccgcgt gttccggctg tcagcgaggg 2100

ggcgcccggt tctttttgtc aagaccgacc tgtccggtgc cctgaatgaa ctgcaggacg 2160  
aggcagcgcg gctatcgtgg ctggccacga cgggcgttcc ttgcgcagct gtgctcgacg 2220  
ttgtcactga agcgggaagg gactggctgc tattgggcga agtgccgggg caggatctcc 2280  
tgtcatctca ccttgctcct gccgagaaag tatccatcat ggctgatgca atgcggcggc 2340  
tgcatacgtc tgatccggct acctgcccat tcgaccacca agcgaaacat cgcatcgacg 2400  
gagcacgtac tcggatggaa gccggctctg tcgatcagga tgatctggac gaagagcatc 2460  
aggggctcgc gccagccgaa ctgttcgcca ggctcaaggc gcgcagccc gacggcgagg 2520  
atctcgtcgt gaccatggc gatgcctgct tgccgaatat catggtggaa aatggccgct 2580  
tttctggatt catcgactgt ggcggctgg gtgtggcgga ccgctatcag gacatagcgt 2640  
tggtacccg tgatattgct gaagagcttg gcggcgaatg ggctgaccgc ttcctcgtgc 2700  
tttacggtat cgcgcgtccc gattcgacgc gcacgcctt ctatcgctt cttgacgagt 2760  
tcttctgagc gggactctgg gggtcgaaat gaccgaccaa gcgacgcca acctgccatc 2820  
acgagatttc gattccaccg ccgccttcta tgaaagggtg ggcttcggaa tcgttttccg 2880  
ggacgccggc tggatgatec tccagcgcg ggatctcatg ctggagttct tcgcccacgc 2940  
tagcggcgcg ccggccggcc cgggtgtgaa taccgcacag atgcgtaagg agaaaatacc 3000  
gcatcaggcg ctcttccgct tcctcgtca ctgactcgt gcgctcggc gttcggctgc 3060  
ggcgagcgg atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 3120  
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 3180  
cgttgctggc gtttttccat aggtccgcc cccctgacga gcatcacaaa aatcgacgt 3240  
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt cccctggaa 3300  
gctccctcgt gcgctctcct gttccgacct tgccgcttac cggatacctg tccgccttcc 3360  
tcccttcggg aagcgtggcg ctttctcata gtcacgctg taggtatctc agttcgggtg 3420  
aggctcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 3480  
ccttatccgg taactatcgt cttgagtcca acccggtgaa acacgactta tcgccactgg 3540  
cagcagccac tggtaacagg attagcagag cgaggatgt aggcgggtgct acagagttct 3600  
tgaagtgggt gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 3660  
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 3720  
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 3780  
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 3840  
aagggatttt ggtcatgaga ttatcaaaaa ggtcttcac ctatgcctt ttaaaggccg 3900  
gccgcggccg ccatcggcct tttcttttgc gtttttattt gttaactgtt aattgtcctt 3960

gttcaaggat gctgtctttg acaacagatg ttttcttgcc ttgatgttc agcaggaage 4020  
tcggcgcaaa cgttgattgt ttgtctgcgt agaatcctct gtttgtcata tagcttgtaa 4080  
tcacgacatt gtttcctttc gcttgaggta cagcgaagtg tgagtaagta aaggttacat 4140  
cgtaggacg aagatccatt tttaacacaa ggccagtttt gttcagcggc ttgtatgggc 4200  
cagttaaaga attagaaaca taaccaagca tgtaaatac gttagacgta atgccgtcaa 4260  
tcgtcatttt tgatccgcgg gagtcagtga acaggtagca tttgccgttc attttaaaga 4320  
cgttcgcggc ttcaatttca tctgttactg tgtagatgc aatcagcggc ttcacactt 4380  
ttttcagtgt gtaatcatcg tttagctcaa tcataccgag agcgccgttt gctaactcag 4440  
ccgtgcgttt tttatcgctt tgcagaagtt ttgactttc ttgacggaag aatgatgtgc 4500  
ttttgccata gtatgctttg ttaaataaag attcttcgcc ttggtagcca tcttcagttc 4560  
cagtgtttgc ttcaaatac aagtatttgt ggcctttatc ttctacgtag tgaggatctc 4620  
tcagcgtatg gttgtgcct gagctgtagt tgccttcac gatgaactgc tgtacatttt 4680  
gatacgtttt tccgtcaccg tcaaagattg atttataatc ctctacaccg ttgatgttca 4740  
aagagctgtc tgatgctgat acgttaactt gtgcagttgt cagtgtttgt ttgccgtaat 4800  
gtttaccgga gaaatcagtg tagaataaac ggatttttcc gtcagatgta aatgtggctg 4860  
aacctgacca ttcttggtt tggcttttta ggatagaatc atttgcatcg aatttgctgc 4920  
tgtctttaa gacgcggcca gcgtttttcc agctgtcaat agaagtttcg ccgacttttt 4980  
gatagaacat gtaaactgat gtgtcatccg catttttagg atctccggct aatgcaaaga 5040  
cgatgtggta gccgtgatag tttgcgacag tgccgtcagc gttttgtaat ggccagctgt 5100  
cccaaactgc caggcctttt gcagaagaga tatttttaat tgtggacgaa tcaaattcag 5160  
aaacttgata tttttcattt ttttgctgtt cagggatttg cagcatatca tggcgtgtaa 5220  
tatgggaaat gccgtatgtt tccttatatg gcttttggtt cgtttctttc gcaaactgtt 5280  
gagttgcgcc tcctgccagc agtgcggtag taaagggtta tactgttgct tgttttgcaa 5340  
actttttgat gttcatcggt catgtctcct tttttatgta ctgtgttagc ggtctgcttc 5400  
ttccagccct cctgtttgaa gatggcaagt tagttacgca caataaaaaa agacctaaaa 5460  
tatgtaaggg gtgacgcaa agtatacact ttgcccttta cacatttttag gtcttgctg 5520  
ctttatcagt aacaaacccg cgcgatttac ttttcgacct cattctatta gactctcggt 5580  
tggattgcaa ctgggtctatt ttctctttt gttgataga aaatcataaa aggatttgca 5640  
gactacgggc ctaaagaact aaaaaatcta tctgtttctt ttcattctct gtatttttta 5700  
tagtttctgt tgcattggga taaagttgcc tttttaatca caattcagaa aatatcataa 5760  
tatctcattt cactaaataa tagtgaacgg caggtatatg tgatgggtta aaaaggatcg 5820

246

gcgcccgctc gatttaaate tcgagaggcc tgacgtcggg 5860

&lt;210&gt; 78

&lt;211&gt; 5860

&lt;212&gt; DNA

&lt;213&gt; Plasmid

&lt;400&gt; 78

cccgtacca cgcgtcccag tggctgagac gcatccgcta aagccccagg aaccctgtgc 60  
agaaagaaaa cactcctctg gctaggtaga cacagtttat aaaggtagag ttgagcgggt 120  
aactgtcagc acgtagatcg aaaggtgcac aaaggtggcc ctggctgtac agaaatatgg 180  
cggttcctcg cttgagagtg cggaacgcat tagaaacgct gctgaacgga tcgttgccac 240  
caagaaggct ggaaatgatg tcgtggttgt ctgctccgca atgggagaca ccacggatga 300  
acttctagaa cttgcagcgg cagtgaatcc cgttccgcca gctcgtgaaa tggatatgct 360  
cctgactgct ggtgagcgta tttctaacgc tctcgtcgcc atggctattg agtcccttgg 420  
cgcagaagcc caatctttca cgggctctca ggtggtgtg ctcaccaccg agcgccacgg 480  
aaacgcacgc attgttgatg tcactccagg tcgtgtgctg gaagcactcg atgagggcaa 540  
gatctgcatt gttgctggtt tccaggggtg taataaagaa acccgcgatg tcaccacgtt 600  
gggtcgtggt gggtctgaca ccactgcagt tgcgttggca gctgctttga acgctgatgt 660  
gtgtgagatt tactcgacg ttgacggtgt gtataccgct gacccgcgca tcgttcctaa 720  
tgacacagaag ctggaaaagc tcagcttcga agaaatgctg gaacttgctg ctggttgctc 780  
caagattttg gtgctgcgca gtgttgaata cgctcgtgca ttcaatgtgc cacttcgctg 840  
acgtcgtct tatagtaatg atccccgcac tttgattgcc ggctctatgg aggatattcc 900  
tgtggaagaa gcagtcctta ccggtgtcgc aaccgacaag tccgaagcca aagtaaccgt 960  
tctgggtatt tccgataagc caggcgaggc tgcaagggtt ttccgtgctg tggctgatgc 1020  
agaaatcaac attgacatgg ttctgcagaa cgtctcttct gtagaagacg gcaccaccga 1080  
catcatcttc acctgccctc gttccgacgg ccgcccgcg atggagatct tgaagaagct 1140  
tcaggttcag ggcaactgga ccaatgtgct ttacgacgac caggtcggca aagtctccct 1200  
cgtgggtgct ggcataagat ctcacccagg tgttaccgca gattcatgg aagctctgcg 1260  
cgatgtcaac gtgaacatcg aattgatttc cactctgag attcgtattt ccgtgctgat 1320  
ccgtgaagat gatctggatg ctgctgcacg tgcattgcat gagcagttcc agctggcgcg 1380  
cgaagacgaa gccgtcgttt atgcaggcac cggacgctaa agttttaag gagtagtttt 1440  
acaatgacca ccacgcagc tggtggtgca accggccagg tcggccaggc tatgcgcacc 1500  
cttttggaag agcgcaattt cccagctgac actgttcggt tctttgcttc cccacgttcc 1560  
gcaggccgta agattgaatt cgtcgacatc gatgctcttc tgcgttaatt aacaattggg 1620

atcctctaga cccgggattt aaatcgctag cgggctgcta aaggaagcgg aacacgtaga 1680  
aagccagtcc gcagaaacgg tgctgacccc ggatgaatgt cagctactgg gctatctgga 1740  
caagggaaaa cgcaagcgca aagagaaagc aggtagcttg cagtgggctt acatggcgat 1800  
agctagactg ggcggtttta tggacagcaa gcgaaccgga attgccagct ggggcgccct 1860  
ctggtaaggt tgggaagccc tgcaaagtaa actggatggc tttcttgccg ccaaggatct 1920  
gatggcgagc gggatcaaga tctgatcaag agacaggatg aggatcggtt cgcattgattg 1980  
aacaagatgg attgcacgca gggtctccgg ccgcttgggt ggagaggcta ttcggctatg 2040  
actgggcaca acagacaatc ggctgctctg atgccgccgt gttccggctg tcagcgcagg 2100  
ggcggccggt tctttttgtc aagaccgacc tgtccgggtg cctgaatgaa ctgcaggacg 2160  
aggcagcgcg gctatcgtgg ctggccacga cggcggttcc ttgcgcagct gtgctcgacg 2220  
ttgtcactga agcgggaagg gactggctgc tattgggcga agtgccgggg caggatctcc 2280  
tgtcatctca ccttgtctct gccgagaaag tatccatcat ggctgatgca atgcggcggc 2340  
tgcatacgct tgatccggct acctgcccac tcgaccacca agcgaaacat cgcattcgagc 2400  
gagcacgtac tcggatggaa gccgggtctg tcgatcagga tgatctggac gaagagcatc 2460  
aggggctcgc gccagccgaa ctgttcgcca ggctcaaggc gcgcatgccc gacggcgagg 2520  
atctcgctgt gacctatggc gatgcctgct tgccgaatat catggtggaa aatggccgct 2580  
ttcttggtat catcgactgt ggccggctgg gtgtggcgga ccgctatcag gacatagcgt 2640  
tggtacccg tgatattgct gaagagcttg gcggcgaatg ggctgaccgc ttcctcgtgc 2700  
tttacggtat cggcgctccc gattcgcagc gcattgcctt ctatcgctt cttgacgagt 2760  
tcttctgagc gggactctgg ggttcgaaat gaccgaccaa gcgacgcca acctgccatc 2820  
acgagatttc gattccaccg ccgccttcta tgaaaggttg ggcttcggaa tcgttttccg 2880  
ggacgccggc tggatgatcc tccagcgcg ggatctcatg ctggagttct tcgcccacgc 2940  
tagcggcgcg ccggccggcc cgggtgtgaaa taccgcacag atgcgtaagg agaaaatacc 3000  
gcatcaggcg ctcttccgct tctcgcctca ctgactcgt gcgctcggtc gttcggctgc 3060  
ggcgagcggc atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 3120  
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 3180  
cggtgctggc gtttttccat aggtccgcc cccctgacga gcatcacaaa aatcgacgct 3240  
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt ccccttgaa 3300  
gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgccttcc 3360  
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg 3420  
aggtcggtcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 3480

ccttatccgg taactatcgt cttgagtcca acccggtaa acacgactta tcgccactgg 3540  
cagcagccac tggtaacagg attagcagag cgagggtatgt aggcggtgct acagagttct 3600  
tgaagtgggtg gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 3660  
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 3720  
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 3780  
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 3840  
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaaggccg 3900  
gccgcggccg ccacggcat tttcttttgc gtttttattt gttaactgtt aattgtcctt 3960  
gttcaaggat gctgtctttg acaacagatg tttcttgcc tttgatgttc agcaggaagc 4020  
tcggcgcaaa cgttgattgt ttgtctgcgt agaatcctct gtttgtcata tagcttgtaa 4080  
tcacgacatt gtttctttc gcttgaggta cagcgaagtg tgagtaagta aaggttacat 4140  
cgttaggatc aagatccatt tttaacacaa ggcagtttt gtccagcggc ttgtatgggc 4200  
cagttaaaga attagaaaca taaccaagca tgtaaatac gttagacgta atgccgtcaa 4260  
tcgtcatttt tgatccgcgg gagtcagtga acaggtacca tttgccgttc attttaaaga 4320  
cgttcgcgcg ttcaatttca tctgttactg tgttagatgc aatcagcggc ttcacactt 4380  
ttttcagtgt gtaatcatcg tttagctcaa tcataccgag agcgccggtt gctaactcag 4440  
ccgtgcgttt tttatcgctt tgcagaagt tttgacttcc ttgacggaag aatgatgtgc 4500  
ttttgccata gtatgctttg ttaaataaag attcttcgcc ttggtagcca tcttcagttc 4560  
cagtgtttgc ttcaaatact aagtatttgt ggcctttatc ttctacgtag tgaggatctc 4620  
tcagcgtagt gttgtcgctt gagctgtagt tgccttcac gatgaactgc tgtacatttt 4680  
gatacgtttt tccgtcacccg tcaaagattg atttataatc ctctacaccg ttgatgttca 4740  
aagagctgtc tgatgctgat acgttaactt gtgcagttgt cagtgtttgt ttgccgtaat 4800  
gtttaccgga gaaatcagt tagaataaac ggatttttcc gtcagatgta aatgtggctg 4860  
aacctgacca ttcttggtt tggcttttta ggatagaatc atttgcacg aatttgctgc 4920  
tgtctttaa gacgcggcca gcgtttttcc agctgtcaat agaagtttcg ccgacttttt 4980  
gatagaacat gtaaategat gtgtcatccg catttttagg atctccggt aatgcaaaga 5040  
cgatgtggta gccgtgatag tttgcgacag tgccgtcagc gttttgtaat ggccagctgt 5100  
cccaaacgtc caggcctttt gcagaagaga tatttttaaat tgtggacgaa tcaaattcag 5160  
aaacttgata tttttcattt ttttgctgtt cagggatttg cagcatatca tggcgtgtaa 5220  
tatgggaaat gccgtatgtt tccttatatg gcttttggtt cgtttcttcc gcaaacgctt 5280  
gagttgcgcc tctgccagc agtgcggtag taaagggttaa tactgttgct tgttttgcaa 5340

249

actttttgat gttcatcggt catgtctcct tttttatgta ctgtgttagc ggtctgcttc 5400  
ttccagccct cctgtttgaa gatggcaagt tagttacgca caataaaaaa agacctaaaa 5460  
tatgtaaggg gtgacgcaa agtatacaact ttgcccttta cacatttttag gtcttgccctg 5520  
ctttatcagt aacaaacccg cgcgatttac ttttcgacct cattctatta gactctcggt 5580  
tggttgcaa ctggtctatt ttcctctttt gtttgataga aaatcataaa aggatttgca 5640  
gactacgggc ctaaagaact aaaaaatcta tctgtttctt ttcattctct gtatttttta 5700  
tagtttctgt tgcattggca taaagttgcc tttttaatca caattcagaa aatatcataa 5760  
tatctcatct cactaaataa tagtgaacgg caggatatatg tgatgggtta aaaaggatcg 5820  
gcggccgctc gatttaaatac tcgagaggcc tgacgtcggg 5860

&lt;210&gt; 79

&lt;211&gt; 8787

&lt;212&gt; DNA

&lt;213&gt; Plasmid

&lt;400&gt; 79

tcgagggaag gtgaatcgaa tttcggggct ttaaagcaaa aatgaacagc ttggtctata 60  
gtggctaggt accctttttt ttttgacac atgtagggtg gccgaaacaa agtaatagga 120  
caacaacgct cgaccgcgat tatttttgga gaatcggtcg ttctccccgg gacgtccac 180  
gacggcgccg accgggcaga ggcaaagccg acagccgtcg catcctaggg agccctttca 240  
tgccctcgtc gccatccacc ccgcccgcgc acaccgcac ccgctgttcc gccctccgag 300  
aggccctcgc caccgcgtg gtggtcgccg acggcgccat gggcaccatg ctccaggccc 360  
agaacccac gctggacgac ttccagcagc tcgaagggtg caacgaggtc ctgaacctca 420  
cccggcccga catcgtccgc tcggtgcacg aggagtactt cggggccggc gtcgactgcy 480  
tcgagaccaa caccttcggc gccaaaccact ccgccctggg cgagtacgac atccccgagc 540  
gcgtccacga actgtccgag gccggcgccc gcgtcgcccg cgaggtcgcc gacgagttcg 600  
gcgcccgcga cggccggcag cgctgggtgc tgggtccat gggccccggc accaagctcc 660  
ccaccctcgg ccacgccccg tacaccgtcc tgcgcgacgc ctaccagcgc aacgccgagg 720  
gactggtcgc gggcggcgcg gacgcactgc tggtgagac cacgcaggac ctgctccaga 780  
ccaaggctc ggtgctcggc gcccggcgcg ccctggacgt cctcggcctc gacctgccgc 840  
tcacgtgtc cgtcacctgc gagaccaccg gcaccatgct gctcggtcg gagatcggcg 900  
ccgctctac cgcgtggaa ccgctcggca tcgacatgat cggcctgaac tgcgccaccg 960  
gccccgccga gatgagcgag cacctgcgtc acctcgcccg gactcccgc atcccgtga 1020  
cctgcatgcc caacgccggt ctgcccgtcc tcggcaagga cggcgccac taccgctga 1080  
ccgccccga gctggccgac gcacacgaga cttctgtgcy cgagtacggc ctgtccctgg 1140

tcggcggctg ctgcggcacc acgcccagac acctgcgcca ggtcgtcgag cgggtccggg	1200
acaccgcccc caccgcacgc gaccgcgccc ccgagcccgg cgccgcctcg ctctaccaga	1260
ccgtgccctt ccgccaggac acctcctacc tggccatcgg cgagcgcacc aacgccaacg	1320
ggtccaagaa gttccgcgag gccatgctgg acggccgctg ggacgactgc gtcgagatgg	1380
cccgcgacca gatccgcgaa ggcgcgcaca tgctcgacct ctgcgtcgac tacgtcggcc	1440
gggacggcgt cgcgcacatg gaggaactgg ccggccgggt cgccaccgce tccacgtgc	1500
cgatcgtcct cgactccacc gaggtcgacg tcatccgggc cggcctggag aagctcggcg	1560
gccgcgcggt gatcaactcg gtcaactacg aggacggcgc cggccccgag tcccggttcg	1620
cccgcgtcac gaagctcgcc cgggagcacg gcgccgcgt gatcgcgctg accatcgacg	1680
aggtgggaca ggcgcgcacc gccgagaaga aggtcgagat cgccgaacgg ctcatcgacg	1740
acctaccgg caactggggc atccacgagt ccgacatcct cgtcgactgc ctgacctca	1800
ccatctgcac cggccaggag gagtcccga aggacggcct ggccaccatc gagggcatcc	1860
gggaactcaa gcggcgccac ccggacgtgc agaccacgt cggcctgtcg aacatctcct	1920
tcggcctcaa cccggccgcc cgcacacctg tcaactccgt cttcctcgac gaatgcgtca	1980
aggccggcct ggactcggcc atcgtgcacg cgagcaagat cctgccgac gcccgcttcg	2040
acgaggagca ggtcaccacc gccctcgact tgatctacga ccgcgcgcg gagggctacg	2100
accccctgca aaagctcatg cagctcttcg agggcgccac cgccaagtcg ctgaaggcct	2160
ccaaggccga ggaactggcc gccctcccgc tggaggagcg cctcaagcgc cgcacatcg	2220
acggcgagaa gaacggcctc gaacaggacc tcgacgaggc cctccgggag cgcccgcccc	2280
tcgagatcgt caacgacacc ctgctcgacg gtatgaaggc cgtcggcgag ctgttcggct	2340
ccggccagat gcagctgccg ttcgtgtccc agtccgcga ggtcatgaag acccggtgg	2400
cccacctgga gccgcacatg gagaagaccg acgacgacgg caagggcacg atcgtgtgg	2460
ccaccgtccg cggcgacgtc cagcacatcg gcaagaacct cgtcgacatc atcctgtcca	2520
acaacggcta caacgtcgtc aacctcggca tcaagcagcc cgtctccgcg atcctggaag	2580
cggccgacga gcaccgggccc gacgtcatcg gcatgtccgg cctcctcgtc aagtccacgg	2640
tgatcatgaa ggagaacctg gaggagctga accagcgcaa gctggccgcc gactaccgg	2700
tcatcctcgg cggcgccgcc ctcaccaggg cctacgtcga acaggacctg cagagatct	2760
acgacggcga ggtccgctac gcccgcgacg ctttcgaggg cctgcgcctc atggacgccc	2820
tcatcggcat caagcgcggc gtgcccggcg ccaagctgcc ggagctgaag cagcgccggg	2880
tcggggcccgc caccgtcgag atcgacgagc gccccgagga agggcacgtc cgctccgacg	2940
tcgccaccga caaccgggtc ccgacccgc cttccgcgg caccgcgtc gtcaagggca	3000



251

tccagctcaa ggagtagcc tcctggctcg acgagggcgc cctcttcaag ggccagtggg 3060  
gcctcaagca ggccccgacc ggcgagggac cctcctacga ggaactggtc gaggccgagg 3120  
gccggccgcg gctgcgcggc ctgctcgacc ggctccagac ggacaacctt ttggaggcgg 3180  
ccgtgggteta cggctacttc ccctgctctt ccaaggacga cgacctgac gtcctcgacg 3240  
acgacggcaa cgaacgcacc cgcttcacct tcccccgcca gcgcgcggc cggcgccctgt 3300  
gcctggccga cttcttccgc ccggaggagt ccggcgagac cgacgtggtc ggcttccagg 3360  
tcgtcacctg cggtccccgc atcggcgagg agacggcccc catgttcgag gccaacgcct 3420  
accgcgacta tctcgagctg cacggcctgt ccgtgcagct cgccgaggcc ctgcgcgagt 3480  
actggcacgc gcgcgtgcgc tcggaactcg gcttcgcccgg ggaggaccgc gccgagatgg 3540  
aggacatgtt cgccctgaag taccgggggtg cccgcttctc cctcggctac ggcgccctgcc 3600  
ccgacctgga ggaccgcgcc aagatcgccg ccctgctgga gcccagagcg atcggcgctc 3660  
acctatccga ggagttccag ctccaccccg agcagtcac cgacgccatc gtcattccacc 3720  
accggagggc caagtacttc aacgcccgtt gagggatatc gtcgacatcg atgctcttct 3780  
gcgttaatta acaattggga tcctctagac ccgggattta aatcgctagc gggctgctaa 3840  
aggaagcgga acacgtagaa agccagtcgc cagaaacggt gctgaccccg gatgaatgtc 3900  
agctactggg ctatctggac aagggaaaac gcaagcgcaa agagaaagca ggtagcttgc 3960  
agtgggctta catggcgata gctagactgg gcggttttat ggacagcaag cgaaccggaa 4020  
ttgccagctg gggcgccctc tggtaagggtt gggaagccct gcaaagtaaa ctggatggct 4080  
ttcttgccgc caaggatctg atggcgagg ggatcaagat ctgatcaaga gacaggatga 4140  
ggatcgtttc gcatgattga acaagatgga ttgcacgcag gttctccggc cgcttgggtg 4200  
gagaggctat tcggctatga ctgggcacaa cagacaatcg gctgctctga tgccgcctg 4260  
ttccggctgt cagcgaggcg gcgcccgggt ctttttgtca agaccgacct gtccggtgcc 4320  
ctgaatgaac tgcaggacga ggcagcgcg ctatcgtggc tggccacgac gggcgttcct 4380  
tgcgacgtg tgctcgacgt tgtcactgaa gcgggaaggg actggctgct attgggcgaa 4440  
gtgccggggc aggatctcct gtcattctac cttgctcctg ccgagaaagt atccatcatg 4500  
gctgatgcaa tgcggcggtt gcatacgctt gatccggcta cctgcccatt cgaccaccaa 4560  
gcgaacatc gcatcgagcg agcacgtact cggatggaag ccggtcttgt cgatcaggat 4620  
gatctggacg aagagcatca ggggctcgcg ccagccgaac tgttcgccag gctcaaggcg 4680  
cgcatgcccg acggcgagga tctcgtcgtg acccatggcg atgcctgctt gccgaatata 4740  
atggtggaaa atggccgctt ttctggattc atcgactgtg gccggctggg tgtggcggac 4800  
cgctatcagg acatagcggt ggctacccgt gatattgctg aagagcttgg cggcgaatgg 4860

gctgaccgct tcctcgtgct ttacggatc gccgctccc attcgcagcg catcgccttc 4920  
tatcgccttc ttgacgagtt cttctgagcg ggactctggg gttcgaaatg accgaccaag 4980  
cgacgcccc cctgccatca cgagatttcg attccaccgc cgccttctat gaaaggttgg 5040  
gcttcggaat cgttttccgg gacgcccggc ggatgatcct ccagcgcggg gatctcatgc 5100  
tggagtcttt cgcacacgct agcggcgcg cgcccgccc ggtgtgaaat accgcacaga 5160  
tgcgtaagga gaaaataccg catcaggcgc tcttcgctt cctcgtcac tgactcgtg 5220  
cgctcggteg ttcggctgcg gcgagcggtc tcagctcact caaaggcgg aatacggtta 5280  
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaggcca gcaaaaggcc 5340  
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgcc ccctgacgag 5400  
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 5460  
caggcgtttc cccctggaag ctccctcgtg cgctctctg ttcgaccct gccgcttacc 5520  
ggatacctgt ccgcctttct cccttcggga agcgtggcg tttctcatag ctacgctgt 5580  
aggtatctca gttcgggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 5640  
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtcaa cccggtgaga 5700  
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 5760  
ggcgggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag aaggacagta 5820  
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 5880  
tccggcaaac aaaccaccgc tggtagcggg ggttttttg tttgcaagca gcagattacg 5940  
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 6000  
tggaacgaaa actcacgtta agggattttg gtcattgagat tatcaaaaag gatcttcacc 6060  
tagatccttt taaaggcccg ccgcgccgc gcaaagtcg gcttcgtgaa aattttcgtg 6120  
ccgctgatt ttcgcctaaa aactttaacg aacgttcgtt ataatggtgt catgaccttc 6180  
acgacgaagt actaaaattg gcccgaaatc tcagctatgg atctctctga tgcgcgctg 6240  
gagtcgcagc cgctcgatgc tgccgtcgat ttaaaaacgg tgatcggatt ttcccgagct 6300  
ctcgatacga cggacgcgcc agcatcacga gactgggcca gtgcccgcag cgacctagaa 6360  
actctcgtgg cggatcttga ggagctggct gacgagctgc gtgctcggcc agcgccagga 6420  
ggacgcacag tagtgaggga tgcaatcagt tgcgcctact gcggtggcct gattcctccc 6480  
cggcctgacc cgcgaggacg gcgcgcaaaa tattgctcag atgcgtgtcg tgccgcagcc 6540  
agccgcgagc gcgccaacaa acgccacgcc gaggagctgg aggcggctag gtcgcaaatg 6600  
gcgctggaag tgcgtccccc gagcgaaatt ttggccatgg tcgtcacaga gctggaagcg 6660  
gcagcgagaa ttatcgcgat cgtggcgggtg cccgcaggca tgacaaacat cgtaaatgcc 6720

gcgtttcgtg tgccgtggcc gcccaggacg tgtcagcgcc gccaccacct gcaccgaatc 6780  
ggcagcagcg tcgcgcgtcg aaaaagcgca caggcggcaa gaagcgataa gctgcacgaa 6840  
tacctgaaaa atgttgaacg ccccgtagcg ggtaactcac agggcgtcgg ctaacccccca 6900  
gtccaaacct gggagaaagc gctcaaaaat gactctagcg gattcacgag acattgacac 6960  
accggccttg aaattttccg ctgatctggt cgacacccat cccgagctcg cgctgcgac 7020  
acgtggcttg acgagcgaag accgccgcga attcctcgct cacctgggca gagaaaattt 7080  
ccagggcagc aagaccgcg acttcgccag cgcttgatc aaagaccccg acacggagaa 7140  
acacagccga agttataccg agttggttca aaatcgctg cccggtgccg gtatgttgct 7200  
ctgacgcacg cgcagcacgc agccgtgctt gtcttgaca ttgatgtgcc gagccaccag 7260  
gccggcggga aaatcgagca cgtaaaccgc gaggtctacg cgattttgga gcgctgggca 7320  
cgctggaaa aagcgccagc ttggatcggc gtgaatccac tgagcgggaa atgccagctc 7380  
atctggctca ttgatccggt gtatgccga gcaggcatga gcagcccgaa tatgcgcctg 7440  
ctggctgcaa cgaccgagga aatgaccgc gttttcgcg ctgaccagc tttttcacat 7500  
aggctgagcc gtggccactg cactctccga cgatcccagc cgtaccgctg gcatgccag 7560  
cacaatcgcg tggatgcct agctgatctt atggaggtg ctgcgatgat ctcaggcaca 7620  
gaaaaacct aaaaacgta tgagcaggag ttttctagcg gacgggcacg tatcgaagcg 7680  
gcaagaaaag ccactgcgga agcaaaagca cttgccacgc ttgaagcaag cctgccgagc 7740  
gccgctgaag cgtctggaga gctgatcgac ggcgtccgtg tcctctggac tgctccaggg 7800  
cgtgccgcc gtgatgagc ggcttttcgc cacgcttga ctgtgggata ccagttaaaa 7860  
gcggctggtg agcgccataa agacaccaag ggtcatcgag cctacgagcg tgcctacac 7920  
gtcgtcagg cggtcggagg aggccgtgag cctgatctgc cgccggactg tgaccgccag 7980  
acggattggc cgcgacgtg gcgcggctac gtcgctaaag gccagccagt cgctcctgct 8040  
cgtcagacag agacgcagag ccagccgagg cgaaaagctc tggccactat gggaagacgt 8100  
ggcggtaaaa aggccgcaga acgctggaaa gacccaaaca gtgagtacgc ccgagcacag 8160  
cgagaaaaac tagctaagtc cagtcaacga caagctagga aagctaaagg aaatcgcttg 8220  
accattgcag gttggtttat gactgttgag ggagagactg gctcgtggcc gacaatcaat 8280  
gaagctatgt ctgaatttag cgtgtcacgt cagaccgtga atagagcact taaggctcgc 8340  
gggcattgaa cttccacgag gacgccgaaa gcttcccagt aaatgtgccg tctcgtagcg 8400  
agaaaacggg tccccgtag ggtctctctc ttggcctcct ttctaggtcg ggctgattgc 8460  
tcttgaagct ctctaggggg gctcacacca taggcagata acgttcccca ccggctcgcc 8520  
tcgtaagcgc acaaggactg ctcccaaaga tcttcaaagc cactgccgcg actgccttcg 8580

cgaagccttg ccccgcgga atttctcca ccgagttcgt gcacaccct atgccaagct	8640
tctttcacc taaattcgag agattggatt cttaccgtgg aaattcttcg caaaaatcgt	8700
cccctgatcg cccttgcgac gttggcgctg gtgccgctgg ttgcgcttgg cttgaccgac	8760
ttgatcagcg gccgctcgat ttaaate	8787